S100 1 and 2

THE OPEN UNIVERSITY

Science Foundation Course Units 1 and 2

Science: Its origins, scales and limitations

Observation and measurement

The Open University

Science Foundation Course Unit 1

SCIENCE: ITS ORIGINS, SCALES AND LIMITATIONS

Prepared by the Science Foundation Course Team

THE OPEN UNIVERSITY PRESS

A NOTE ABOUT AUTHORSHIP OF THIS TEXT

This text is one of a series that, together, constitutes *a component part* of the Science Foundation Course. The other components are a series of television and radio programmes, home experiments and a summer school.

The course has been produced by a team, which accepts responsibility for the course as a whole and for each of its components.

THE SCIENCE FOUNDATION COURSE TEAM

M. J. Pentz (Chairman and General Editor)

F. Aprahamian	(Editor)	J. McCloy	(BBC)
A. Clow	(BBC)	J. E. Pearson	(Editor)
P. A. Crompton	(BBC)	S. P. R. Rose	(Biology)
G. F. Elliott	(Physics)	R. A. Ross	(Chemistry)
G. C. Fletcher	(Physics)	P. J. Smith	(Earth Sciences)
I. G. Gass	(Earth Sciences	F. R. Stannard	(Physics)
L. J. Haynes	(Chemistry)	J. Stevenson	(BBC)
R. R. Hill	(Chemistry)	N. A. Taylor	(BBC)
R. M. Holmes	(Biology)	M. E. Varley	(Biology)
S. W. Hurry	(Biology)	A. J. Walton	(Physics)
D. A. Johnson	(Chemistry)	B. G. Whatley	(BBC)
A. B. Jolly	(BBC)	R. C. L. Wilson	(Earth Sciences)
A. R. Kaye	(Educational Technology)		

The following people acted as consultants for certain components of the course:

J. D. Beckman	B. S. Cox
H. G. Davies	R. J. Knight
D. J. Miller	M. W. Neil
J. R. Ravetz	H. Rose

The Open University Press
Walton Hall Milton Keynes
MK7 6AA

First Published 1970. Reprinted 1971, 1972, 1973, 1976
Copyright © 1970 The Open University

All rights reserved.
No part of this work may be reproduced in any form, by mimeograph or by any other means, without permission in writing from the publishers.

Designed by the Media Development Group of the Open University

Printed in Great Britain by
EYRE AND SPOTTISWOODE LIMITED
AT GROSVENOR PRESS PORTSMOUTH

SBN 335 02000 3

Open University courses provide a method of study for independent learners through an integrated teaching system, including textual material, radio and television programmes and short residential courses. This text is one of a series that makes up the correspondence element of the Science Foundation Course.

For general availability of supporting material referred to in this text, please write to the Director of Marketing, The Open University, P.O. Box 81, Milton Keynes, MK7 6AT.

Further information on Open University courses may be obtained from the Admissions Office, The Open University, P.O. Box 48, Milton Keynes, MK7 6AB.

Contents

	List of Scientific Terms, Concepts and Principles	5
	Conceptual Diagram	6
	Objectives	8
1.1	**Science, Technology and Society**	9
1.2	**Is Human 'Science' Unique?**	13
1.2.1	The senses and their development	13
1.2.2	The development of the brain	15
1.2.3	Memory and learning	16
1.2.4	'Exploratory' behaviour	17
1.2.5	Generalization	18
1.2.6	Tools	18
1.2.7	Speech	19
1.3	**The Birth of Science**	22
1.3.1	The Renaissance	23
1.4	**Scientific Knowledge and Craft Knowledge; an experiment**	25
1.4.1	Galileo's inclined-plane experiment	25
1.4.2	How Galileo described the first part of his experiment	26
1.4.3	The first experiment: what you have to do	27
1.4.4	Introduction to the second part of Galileo's experiment	29
1.4.5	The second experiment: what you have to do	30
1.4.6	A quick test of the extrapolation from inclined plane to free fall	32
1.4.7	Some remarks about errors	35
1.5	**The Nature of Scientific Knowledge**	36
1.6	**The Creative Work of Science**	38
1.6.1	Limitations on research	39
1.7	**Establishing Scientific Results**	40
1.7.1	The design of experiments	41
1.7.2	Models	42
1.7.3	Hypotheses	43
1.7.4	Deductive and inductive inference	44
1.8	**The Achievement of Scientific Knowledge**	45
1.8.1	Facts and laws	46
1.9	**Summary**	48
	Book List	48
	Appendix 1 (Red) Some Examples of Probability	49
	Appendix 2 (Black) Theory of the Galileo Experiment	50
	Appendix 3 (White) Glossary	53
	Self-Assessment Questions	54
	Self-Assessment Answers and Comments	60

Explanatory note

In this and subsequent units of this Course, the preliminary pages will include:

(1) a list of scientific terms, concepts, and principles used in the unit;
(2) a conceptual diagram of the main ideas discussed in the unit;
(3) a statement of learning objectives for the unit.

We hope these items will help you in planning your work, in relating units one to another, and in revising during the course of the year.

The *list of objectives* is a guideline to what we hope you will learn from each unit, and an indication of what you may expect in assessment and examination material. You will find that the self-assessment questions relate to and exemplify the objectives.

The *list of scientific terms, concepts, and principles* was originally used to help prepare the Course. We believe that they can also help you in your studies and revision. The tables indicate:

(i) what you are assumed to know before you start work on the unit (columns 1 and 2);
(ii) the main terms you are expected to learn, and where they are discussed in the unit (column 3);
(iii) how the unit is related to subsequent units (column 4).

The *conceptual diagram* is a simplified interpretation of the relationships between the main ideas discussed in the unit. Each diagram should help you to visualize the whole subject matter of the unit, and to stimulate you to produce your own diagram should you think it necessary.

In the main text of the unit you will occasionally find a question that is answered in the margin. The first example of such a question is on page 13 of Unit 1. If you wish to try to answer the question without seeing the answer, just cover the margin with a suitable piece of card or paper. This is really a matter of personal preference; there is no evidence that you are likely to learn better or worse if the answer is 'hidden'.

Table A

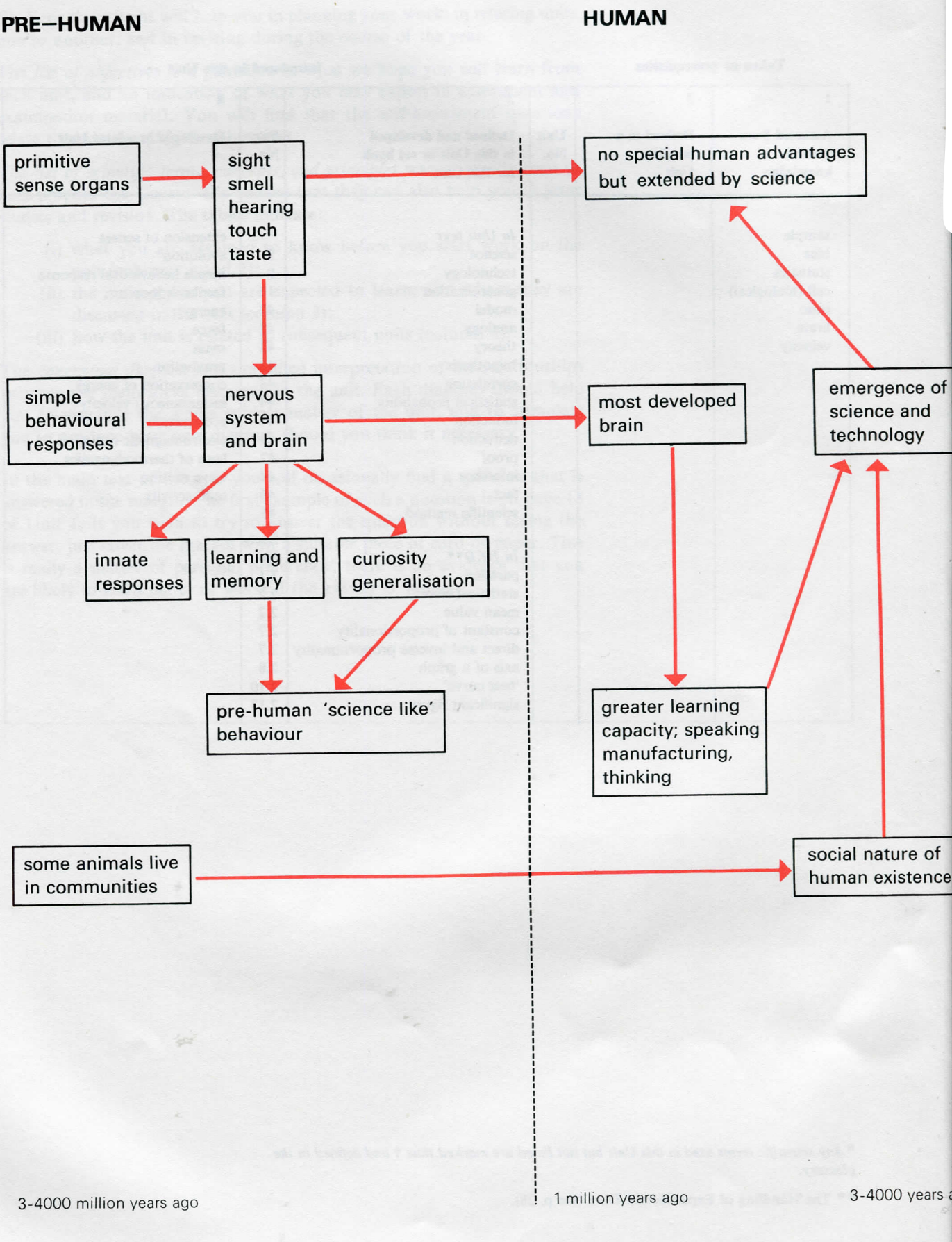
A List of Scientific Terms, Concepts and Principles used in Unit 1*

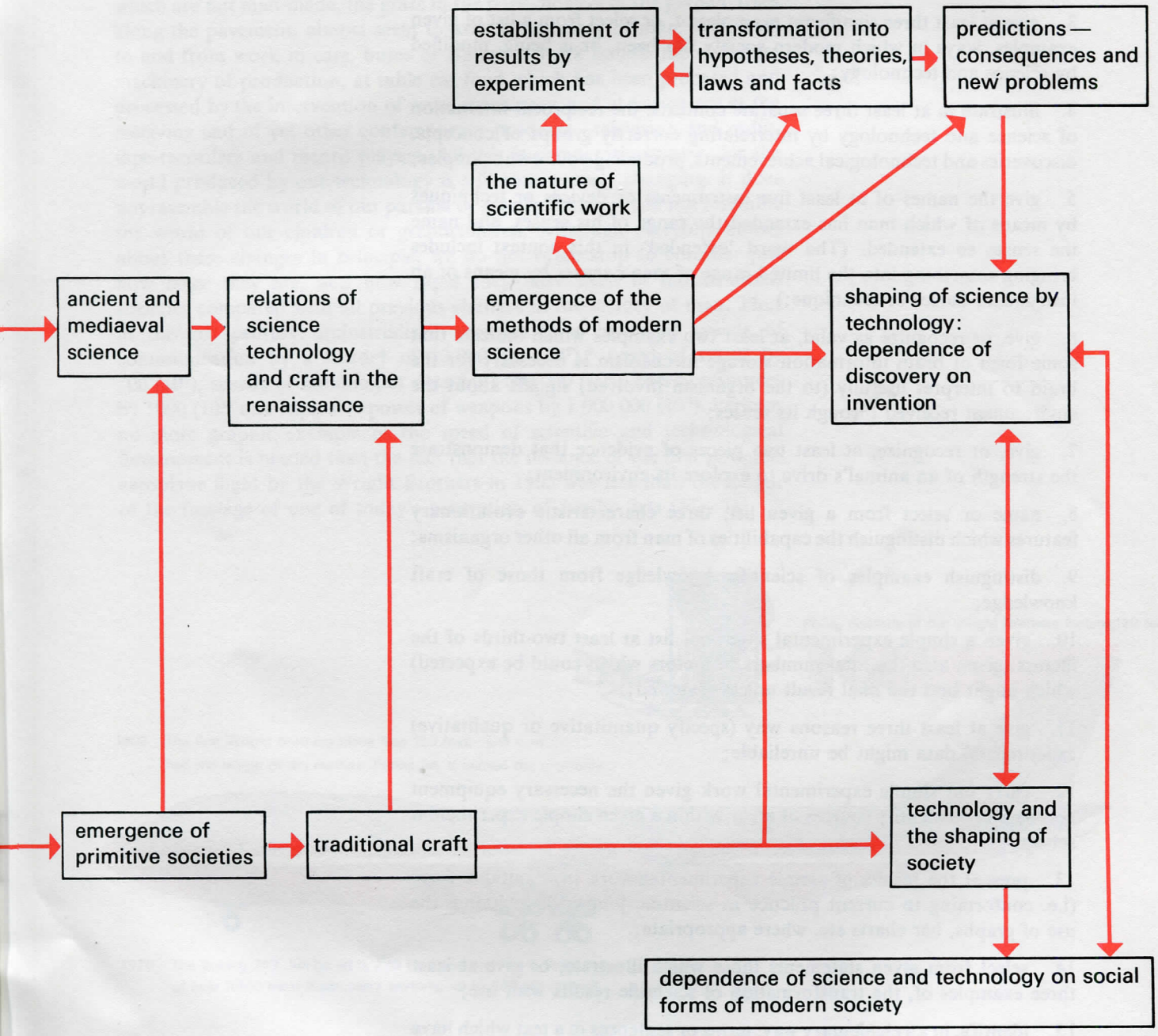
Taken as prerequisites			Introduced in this Unit			
1 **Assumed from general knowledge**	**2** **Defined in a previous Unit**	**Unit No.**	**3** **Defined and developed in this Unit or set book for this Unit**	**Page No.**	**4** **Developed in a later Unit**	**Unit No.**
sample			*In Unit text*		extension of senses	2
bias			science	9	evolution	21
statistics			technology	9	innate behavioural response	20
cell (biological)			generalization	16	feedback loop	16
ratio			model	42	energy	4
brain			analogy	43	force	4
velocity			theory	42	mass	3
			hypothesis	43	gravitation	4
			correlation	43	conservation of energy	4
			statistical probability	44	instantaneous velocity	3
			induction	44	acceleration	3
			deduction	44	electromagnetic radiation	2, 28
			proof	42	laws of thermodynamics	5
			inference	44	wave motion	2, 22
			fact	46	radioactivity	6, 31
			scientific method	45		
				Sect. No.		
			*In HED***			
			parallax	1.1		
			statistical error	2.2		
			mean value	2.2		
			constant of proportionality	2.7		
			direct and inverse proportionality	2.7		
			axis of a graph	2.8		
			'best curve'	2.10		
			significant figure	2.13		

**Any scientific terms used in this Unit but not listed are marked thus † and defined in the glossary.*

** The Handling of Experimental Data (see p. 28).

MAN'S PROGRESS IN SCIENCE & TECHNOLOGY




establishment of results by experiment
transformation into hypotheses, theories, laws and facts
predictions — consequences and new problems
the nature of scientific work
ancient and mediaeval science
relations of science technology and craft in the renaissance
emergence of the methods of modern science
shaping of science by technology: dependence of discovery on invention
emergence of primitive societies
traditional craft
technology and the shaping of society
dependence of science and technology on social forms of modern society
300 years ago
today

Objectives

When you have completed this course unit you should be able to:

1. define in your own words or recognize valid definitions of the terms, concepts and principles listed in column 3 of Table A (p. 5), and recognize situations implicitly involving the ideas expressed by these terms, concepts and principles;

2. give six examples of the orders of magnitude by which man's technological capabilities have increased over the last century;

3. give at least three significant examples of, or select from a list of given examples, ways in which modern society has been, or is being, modified by science and technology;

4. illustrate in at least three separate contexts, the reciprocal interaction of science and technology by interrelating correctly groups of concepts, discoveries and technological achievements, producing your own examples;

5. give the names of at least five instruments or devices or techniques by means of which man has extended the range of his senses, and name the senses so extended. (The word 'extended' in this context includes bringing something into the limited range of man's senses by means of an instrument, device or technique;)

6. give, or recognize as valid, at least two examples which indicate that some form of brain information storage mechanism is necessary for the brain to interpret usefully (to the organism involved) signals about the environment received through its senses;

7. give, or recognize, at least two pieces of evidence that demonstrate the strength of an animal's drive to explore its environment;

8. name or select from a given list, three characteristic evolutionary features which distinguish the capabilities of man from all other organisms;

9. distinguish examples of scientific knowledge from those of craft knowledge;

10. given a simple experimental situation, list at least two-thirds of the factors (given also the total numbers of factors which could be expected) which might bias the final result unless reviewed;

11. give at least three reasons why (specify quantitative or qualitative) experimental data might be unreliable;

12. carry out simple experimental work given the necessary equipment and be able to identify sources of error within a given simple experimental set-up;

13. present the results of simple experimental work in a suitable form (i.e. conforming to current practice in scientific journals), including the use of graphs, bar charts etc. where appropriate;

14. select from given statements those which illustrate, or give at least three examples of, the transformation of scientific results with use;

15. identify, in an elementary way, terms or sentences in a text which have functions or natures or roles which illustrate the meanings of such concepts as: making and testing hypotheses; induction. deduction; argument by analogy; facts; principles;

1.1 Science, technology and society

We live in a highly complex industrial society. For those of us—80 per cent of the population of Britain—who live in towns and suburbs, much of our lives is passed in an environment which is man-made. Even those bits which are not man-made, the grass in the park, flowers in the garden, trees along the pavement, almost seem to exist by courtesy of man. We travel to and from work in cars, buses or trains, at work handle the elaborate machinery of production, at table eat food which has been prepared and processed by the intervention of more machinery, and at leisure are at the receiving end of yet other contemporary products—television and radio, tape-recorders and record players. All this is comparatively new, for the world produced by our technology is still continuously changing. It does not resemble the world of our parents or grandparents; it will not resemble the world of our children or grandchildren. Although everyone knows about these changes in principle, we do not often stop to consider just how large they are, and how rapid they have been in industrialized societies compared with all previous changes in the history of man. Thus in the last century, industrialized societies have increased speeds of communication by a factor† of 10 000 000 (10^7),* speed of travel by 100 (10^2), speeds of handling of data by 1 000 000 (10^6), energy resources by 1000 (10^3) and explosive power of weapons by 1 000 000 (10^6). Perhaps no more graphic example of the speed of scientific and technological development is needed than the fact that the distance of the first powered aeroplane flight by the Wright Brothers in 1903 was less than the length of the fuselage of one of today's generation of Jumbo Jets (Fig. 1).

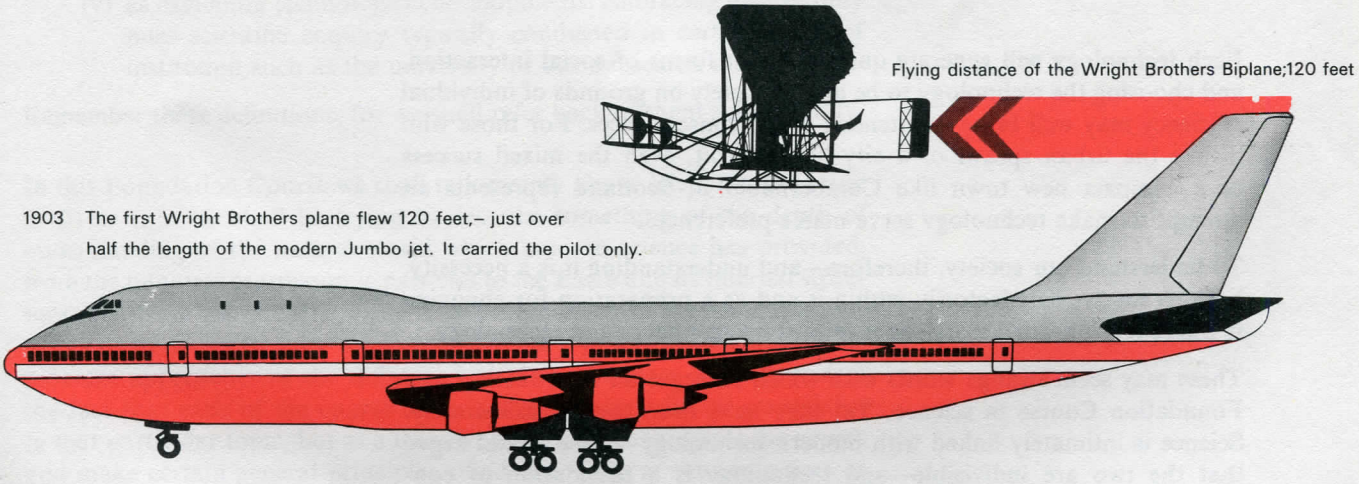

Flying distance of the Wright Brothers Biplane;120 feet
1903 The first Wright Brothers plane flew 120 feet,– just over half the length of the modern Jumbo jet. It carried the pilot only.
1970 The Boeing 747 Jumbo jet is 230 feet long and has a flying range of over 3500 miles.It can carry anything up to 490 passengers

Figure 1

† *The meaning of terms marked thus* † *will be found in the glossary* (*p.* 53).

* *If you are unfamiliar with this way of writing numbers, an explanation of it will be found in the book* Mathematics for the Foundation Course in Science, *section* 1*A. This book will be referred to as* MAFS.

Because we take for granted so much of modern technology, it is difficult now to imagine what life was like only a few generations ago, before the development of today's 'technological revolution'. It is difficult to imagine how we would live if we were transported to a desert island without the products of modern technology and left to fend for ourselves. Present-day societies are held together by a set of technologies which have to some extent determined their shape—and contributed to their problems. So, too, to some extent, were earlier societies. To take but a simple example, whether a village is watered from a central well or from a drain which runs down the sides of its roads substantially determines whether the village is built in a compact central cluster or whether it is spread out in long ribbons. Whether a city is served by an effective public transport system or depends predominantly on the private car will help determine whether neighbourhoods within it are close-packed or diffuse and linked by motorways spreading amorphously into the surrounding countryside (Figs. 2a and b).

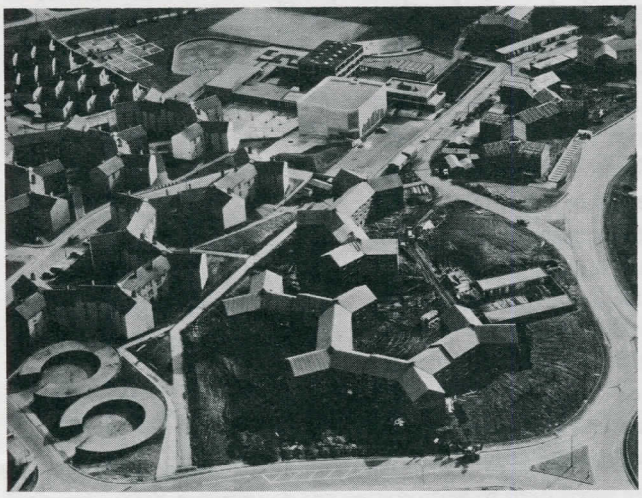


Figure 2a *Cumbernauld.*

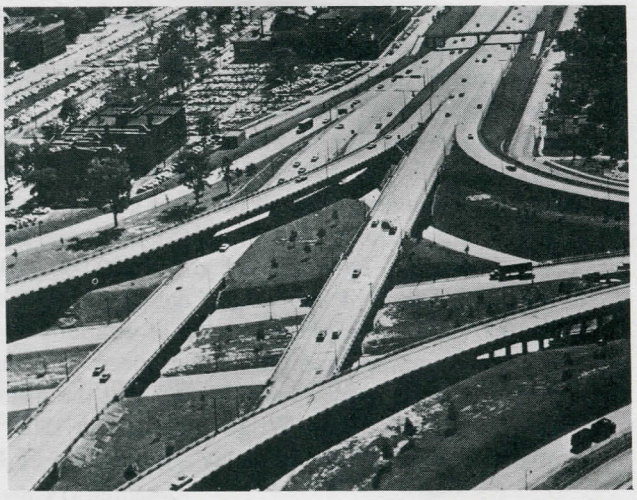


Figure 2b *Detroit.*

Each technology will generate quite different forms of social interaction, and choosing the technology to be applied solely on grounds of individual efficiency may well bring unintended social consequences. For those who dislike the urban sprawl of a city like Detroit, even the mixed success of a compact new town like Cumbernauld in Scotland represents an attempt to make technology serve man's preferences.

To understand our society, therefore—and understanding it is a necessity, both to survive satisfactorily within it and as a preparation for changing it where it is unsatisfactory—it is important to understand technology.

These may seem strange words with which to begin an introduction to a Foundation Course in science. Yet they have been chosen deliberately. Science is intimately linked with modern technology—some would argue that the two are indivisible—and technology is a foundation of contemporary society. The activities of scientists are an attempt to explain the natural and human world, those of technologists to use these explanations so as to manipulate this world, to use its properties in order to build new objects, machines or devices. But there is a continuity between these activities. Science traditionally may be taken to mean the advancement of our understanding of the way in which the observable world works, the development of logical, integrated and self-consistent descriptions of how and why such and such individual happenings occur, how apples fall from trees, how they appear to be coloured red and green, why they are good to eat, irrespective of the immediate usefulness of these statements. That is to say, a derivation of the laws of gravity or of optics, chemical

analysis of the constituents of the apple, or a knowledge of the physiology of digestion, is not supposed to make us immediately richer, stronger, or more able to control nature. On the other hand, the invention of the telescope, of techniques for the cooking, canning, bottling or preserving of apples, or of medicines to soothe the stomach pains we will get if we eat too many of them, are seen as technologies—they do not add to our understanding of the working of the laws of nature, but they add to our control over the world around us.

In fact, such a distinction cannot be long maintained. Telescopes were the product of a study of the properties of lenses by a Dutch spectacle-maker; but in their turn they were developed by the Italian, Galileo, for sale to the merchant traders of Venice as well as a means to study the planets. This study, in its turn, led to the discovery of the moons of Jupiter, and ultimately to evidence that the Earth revolved around the Sun and to the development of a theory of gravitation. Even the medical development of the use of magnesium hydroxide to alleviate the discomfort due to stomach acidity also led on to evidence that the stomach juices were acidic, and that their acidity was regulated by the body, ensuring efficient digestion.

So science and technology must be seen as interdependent activities; discovery precedes invention, and invention in turn makes possible discovery—at least in our contemporary society. These links between science and technology, between the methods of science and the results of science, make it difficult to agree on a single precise definition of either 'science' or 'technology'.

There are indeed at least five current uses of the word science:

(i) as the attempt to discover and explain the workings of the world of nature;
(ii) as the application of certain rules of procedure and enquiry;
(iii) as the social institutions within which these activities are carried out;
(iv) as including the whole field of research and development, that is, both science and technology as distinguished above;
(v) as excluding technological developments, embracing instead only pure scientific enquiry typically conducted in certain types of institution such as the university or basic research institute.

Remember these definitions, for we shall refer back to them subsequently.

In this Foundation Course we shall talk mainly about science in the sense of (i) or (ii). We shall be trying to describe something of the depth of understanding of the natural world which modern science has provided, from the minutest of subatomic particles to the Earth and its internal structure, and from the simplest virus to the mechanism of the human brain. This course is written by an interdisciplinary group of scientists. The different approaches of the different branches of science are reflected in the fact that each of the course units is written primarily by a specialist in that particular topic, but as a team we have tried to integrate the course and make certain general principles stand out. Thus at the same time as presenting you with the *facts* and *theories* of contemporary science, we are also trying to show you something about how a scientist goes about his business; how he makes experiments, how he happens upon theories and explanations, what he accepts as proof. We take account also of the technological and hence social application of the results of scientific research. The tremendous scientific achievements that we describe have not been made in some sort of social vacuum, unrelated to the developments in the surrounding society, and the scientist in his white lab coat is not a man apart from society. In this sense the foundations of science are also the foundations of technology—and, consequently of society. It is in making these links clear, and in attempting both to present the

panorama of modern scientific thought in a way that is broad and deep, and to show the relationships between sciences sometimes regarded as separate, like, for example, geophysics and biochemistry, that the authors of this course think that it differs from any conventional university courses in the sciences, and attempts something both new and, we think, important.

Only your experience through this course will establish how far we have succeeded.

As a check, write down now the reasons why you have elected to take the Foundation Course in science and three things or problems that you hope you will be able to understand at the end of it. At the end of the year, you will be able to assess how well you—and the course—have done.

Reasons

1

2

3

Things or problems to be understood

1

2

3

1.2 Is human 'science' unique?

It is usual to set the emergence of 'modern science'* as occurring in Europe, in the sixteenth century and the development of science as proceeding with ever-increasing speed thereafter, particularly following the Industrial Revolution of the nineteenth century, in Britain, France and Germany, and rather later in the United States, Russia and Japan. But the activities described as science are the natural continuation of processes that had, indeed, been going on in earlier times in human societies. What is more, they are an extension of developments which we can trace not only in humans but even in the most primitive animal.

1.2.1 The senses and their development

The evolutionary path which has led to man and his society has been that of ever-increasing capacity to sense and react to the external world.

Man is commonly described as having five senses. Can you name them?

Sight, hearing, touch, taste, smell.

With these senses, primitive man must have explored the surrounding world, learned what was and what was not good to eat, what was to be feared and avoided as dangerous and what was not. Much of modern science has depended on the process of extending the range and precision of these senses, particularly those of sight, hearing and smell, making possible the detection of objects at a distance or with a degree of accuracy and clarity not otherwise possible, or of substances which would otherwise be outside the limits of detectability.

Name some scientific devices which extend the range of each of the three senses, sight, sound or smell. (Try to name at least one instrument which corresponds to each sense.)

Sight—telescope, microscope.
Sound—stethoscope, telephone.
Smell—breathalyser, smoke detector.

Maybe you got the first two right and not the third. The sense of smell depends on the presence in the nose of special detectors which respond, in ways we will describe in Unit 18, to small quantities of individual chemical substances in the air breathed in. The breathalyser, in measuring the amount of alcohol in breath, performs a similar job.

But the senses did not arise with man alone. More primitive animals possess them, often more fully developed than in man.

Name some animals which have better sight, hearing or smell than man.

You might have chosen from amongst many. For example, the ranges of sound that can be heard by the bat or the dog exceed that of man. The eyes of the owl are adapted to see quite well in a degree of darkness in

* *If you would like to read more about this, see the relevant parts of Ravetz and Bernal (Vol. 2, Chapters 7, 8 and 9). We also suggest that you read Chapters 1 to 6 of Rose and Rose over the course of the year. We shall refer to these Chapters again in Units 33 and 34. The full details of these books are given in the Book List on p. 48.*

which man cannot see at all. Not only can dogs respond to smells undetectable to the human nose, but many insects, such as bees and moths, can detect chemicals (such as the sex attractants, produced by some female moths) at concentrations so low as to be hard to detect even with the most sensitive instruments yet devised.

Man uses his senses to acquire information about his environment, particularly as to food and danger, a characteristic he shares in common with all other animals. Even the simplest one-celled creatures can react to their environment. (If you do not know in general terms that all living organisms are built up from units known as *cells*, there is no need to worry unduly here; you will learn about them in Unit 14.) Put an obstacle in the way even of a minute one-celled organism and it will move around it; put a drop of the corrosive substance sulphuric acid in the middle of a dish containing the creatures and they will move away from it. Similarly such organisms will move towards concentrations of foodstuffs or towards or away from a source of light. Such reactions to obstacles and to food are the primitive versions of the senses known in man as touch and taste or smell. A sense of temperature, that is of hot and cold, seems likely to have evolved at about the same time as these other fundamental senses. But perhaps the most significant evolutionary development was that of vision which, in evolutionary terms, probably developed quite early. The simplest present-day animals to possess a rudimentary form of vision are the little pond creatures known as flat worms. Visual patterns can be interpreted not just in terms of generalized responses at the surface of the organism, but in terms of the world of distant objects. But for this to be done, the eye must be accompanied by a *brain* capable of interpreting a complex set of stimuli. Conversely, since no other sensory organ in an animal provides such a wealth of information it is likely that the development of the human brain was linked to that of the eye.

What may have happened is that the primitive light sense shown by the one-celled animal became specialized in more complex organisms. Certain regions and groups of cells became especially sensitive to light. The visual sense probably developed from a response to moving shadows on the surface of the skin, which would have given warning of nearby danger, to recognition of patterns, when eyes developed optical systems. The stages seem to have been first a concentration of specially light-sensitive cells located at certain regions of the organism, and then 'eye-pits' with the light sensitive cells with pigmented cells round them forming the bottom of the pits. These pits gradually deepened in more advanced organisms, serving to increase the contrast of shadows in the light sensitive region by shielding them from surrounding strong light. The lens of the eye most probably started as a transparent window, protecting the eye-pits from being blocked by small particles floating in the sea in which primitive organisms lived.The protective windows may have gradually thickened at their centres, for this would have increased the intensity of light on the sensitive cells, until the central thickening became a *lens*, producing an image-forming eye, which could present optical patterns to the primitive touch-sensitive nervous system (you will learn more of the properties of such lenses in Unit 2 and of the evolution of organisms in Unit 21).

From the point in time when eyes (and ears) developed, together with the brains to process the information they provided, one road of evolutionary success lay in the development of more accurate senses and more powerful brains. Sight and hearing enabled information arriving at the organism from a distance to be processed, analysed and acted upon, giving warning of the future, enabling predators to see and hear their prey—and prey to see and hear their predators.

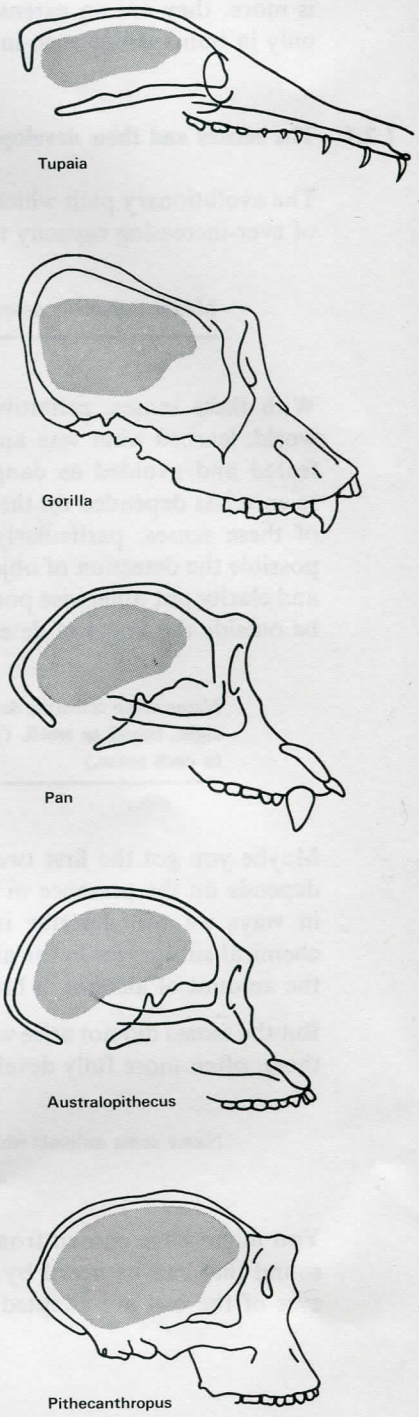


Figure 3

1.2.2 The development of the brain

The development of the senses has not been the only pathway to success in the evolution of life—plants, for instance, have developed along a different route—but for the animal kingdom it has been amongst the major ones. You will be learning in Unit 2 how one section of science is based on this development of the senses, and their extension by means of instruments and other sophisticated measuring and transforming devices that enable men to see the otherwise unseeable, hear the unhearable, smell the unsmellable, and thus enlarge their *understanding of* (i.e. science) and *power over* (i.e. technology) the external world. These extensions are in a way analogous to the extensions of our senses provided by the arts, for instance by pictures or music. These, too, in a way extend our knowledge of the world external to man, and hence our understanding of man's relationships with the world.

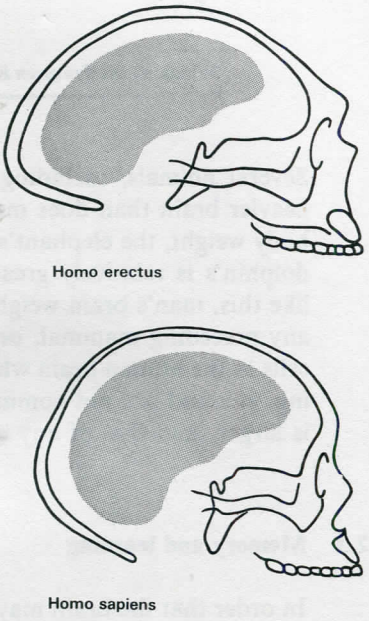


Figure 3 (*cont.*)

With increased powers of sensing the environment come increased possibilities of doing something about it. The fully developed brain not only interprets the data received by way of the senses; it also decides on action to be taken and then signals to other parts of the body to make the appropriate responses. Simple signalling systems exist even in primitive animals. Even in one-celled animals there are methods for one end of the animal to keep the other end posted as to events in its environment. In many-celled animals the signalling system became more elaborate and specialized signalling cells, nerve cells, evolved. Parallel with the development of eyes, groups of nerve cells accumulated towards the front end (the end with the eyes) of the creature to form a simple brain, which became responsible for receiving the information from the eyes and other senses and processing it. The evolutionary pathway that has led to man has been one of steadily increasing brain complexity, size and number of component units (cells). Some of this development can be seen in Figure 3.

Table 1

Brain and body weights and brain cell numbers in some mammals

Animal	Brain weight (kg)	Body weight (kg)	Approx. number of brain cells 'available for learning' (in millions)	Brain/Body weight ratio	Million brain cells per kg body weight
Rat	0.002	0.300	20	0.007	67
Elephant	6.00	7 000	18 000	0.001	3
Dolphin	1.75	150	10 000	0.012	67
Gorilla	0.60	250	3 600	0.002	14
Chimpanzee	0.40	45	3 400	0.009	75
Baboon	0.20	20	2 100	0.010	105
Macaque monkey	0.10	10	1 200	0.010	120
Homo Erectus	0.90	50	not known	0.018	—
Homo Sapiens (man)	1.30	60	8 500	0.022	146

Table 1 compares a number of different mammals, including various apes and one of the immediate evolutionary ancestors of man, known as *Homo erectus*, as well as man, *Homo sapiens*, himself. It shows brain and body weights and the number of brain cells involved in learning processes, and then a calculation of two ratios, those of brain weight to body weight, and of number of brain cells per kilogram of body weight. (The figures for the extinct human ancestors are estimates based on measurements of skull

size.) Study the table carefully, particularly if you are not familiar with extracting information from such tables, then answer the question:

Which of the features listed are greatest in man?

The ratio of the brain cell number to body weight, and of brain weight to body weight.

Several animals, including both the elephant and the dolphin, have a heavier brain than does man; but compared on the basis of a 'standard' body weight, the elephant's brain is smaller even than the rat's, whilst the dolphin's is relatively greater than that of some of the apes. Expressed like this, man's brain weight and brain cell number is higher than that of any preceding mammal, or other animal yet examined. The number of cells in the human brain which are 'free' for thinking, learning, remembering, etc. and are not committed to particular stereotyped body functions is larger than that of any other animal.

1.2.3 Memory and learning

In order that the brain may signal an appropriate response to information received about a situation, it must be able to *compare* this present information with other material which in the past resembled it. Is the noise arriving at the ears that of an enemy or possibly of prey? Is the shape and pattern observed by the eyes that of friend or foe? To answer these questions it is necessary to know whether similar noises and shapes in the past have been those of predator or prey, friend or foe, and to compare the present ones with the past. This comparative procedure is that of memory, and even *apparently* quite simple animals show memory and can learn. Insects such as bees can remember quite complex 'maps' of how to get from their hives to food and back again. Snails and some insects show certain types of learning. Thus if a snail receives an electric shock every time it extends its eye stalk, it will learn not to do so; if an octopus receives a crab for food every time a white square is shown at the same time, but an electric shock when the crab is presented at the same time as a black square, the octopus will learn not to approach the crab when the black square is shown. Even cockroaches can be trained to fold their legs so as to avoid getting an electric shock. More complex organisms, such as pigeons and other birds, can be taught to count (up to 4 or 5 at any rate) for reward of food, whilst rats can be trained to traverse quite complex mazes, press levers for food and distinguish between circles, squares and other shapes. Any such relatively crude tests of memory and learning that can be studied by the scientist in his laboratory are elementary feats by comparison with those achieved by animals in their everyday lives. An insect-eating bird needs to remember not only that 'flies taste good' but also that 'yellow and black striped flies taste bad'. This is the procedure of making *generalizations* about the world; 'if one (or two or three) yellow and black fly tastes nasty and is dangerous, then *all* may well do'. Thus the animal develops a set of behaviour patterns (or rules) for ordering and arranging the universe around it. (However, such rules can lead to mistakes being made. Some good-tasting insects have also developed yellow and black striped colouring and so survive better because they appear to be dangerous.)

Some of these rules of procedure have apparently become, during the course of evolution, *innate*, that is inherited or inborn. The baby birds of certain species open their mouths when they see their parent, and thus can receive food; they will do this, even when hatched and reared in isolation and fed by hand, whenever they are shown even a crude 'beak' shape, provided it is coloured appropriately. Other species respond to the shadow of a hawk-like shape by cringing or seeking shelter, even when

they have been reared from birth under conditions in which they could not possibly ever have seen a bird of prey. Such behavioural responses are thus built into the animal (you will read more about them in Unit 20).

1.2.4 'Exploratory' behaviour

But whilst these responses are perhaps so important for survival that they seem inflexible to the outside observer, some species, particularly along the evolutionary path that has led to man, also show a considerable degree of flexibility in response, enabling the animal to adapt itself and learn to survive under changing conditions in the external world. Perhaps this flexibility is best shown by the powerful 'drive' which exists in some animals towards 'exploratory' behaviour. For instance, if a rat is made hungry and thirsty by depriving it of food and water for 24 hours, and then both are put into its cage, it will immediately both eat and drink. But if instead of the food and water being placed in the rat's own cage, it is taken out of its cage and placed in a fresh one where there is food and water, it will *not* eat and drink immediately. Instead, it will roam up and down the new cage, sniffing at it, apparently examining it and exploring it from many angles. Only later will it turn to the food and water, and satisfy hunger and thirst. The rat behaves as if its 'drive' to explore a new and unfamiliar situation is even more powerful than that to eat and drink.

Although one must beware of making analogies between the behaviour of rats and that of humans (such a tendency verges upon *anthropomorphism* —investing a non-human thing with human characteristics, a tendency often displayed by doting dog-owners), one can cautiously ask the question:

What behaviour pattern in humans would you say would be similar to that shown by the rat?

Curiosity.

Such 'curiosity-like' behaviour is shown in species of the mammalian evolutionary line which leads to man.

An even more striking example is that rats will behave as if they are adopting and testing simple strategies and hypotheses. In one group of experiments, made in California in 1935, rats were shown a series of closed doors under each of which a light could be switched on. Only one light was switched on at a time, and all the doors except one were locked. The rat had to jump at a door: if it was unlocked he could get through and was rewarded with food. The only clue as to which was the unlocked door was the position of the light. It could be the door above the light, or the one to the right, or in any other relation. In this situation the behaviour of the rats could only be interpreted (by humans) as if the rats were testing their own 'hypotheses', for instance, the hypothesis that the unlocked door was the one to the right of the light whenever the light appeared. If after several trials they failed, they would try another hypothesis until they hit on the right one. The point of the experiment was that the rats did not try just any combination of doors and lights, they worked through and 'tested' particular 'hypotheses' quite systematically.

Even more clearly explorative behaviour is shown in the animal species closest to man, such as the apes. During the First World War, a famous series of experiments was conducted by the German psychologist Wolfgang Köhler in which he placed a chimpanzee in a cage with a bunch of bananas just outside. The animal soon learned to reach outside the cage for the bunch. Köhler then put the bunch of bananas further from the cage, so that it could not be reached by the monkey simply stretching out its hand. But he put a stick in the cage. After failing to reach the bananas with its

hands, the animal eventually picked up the stick, poked it through the bars of the cage and hooked the bananas with the end of it. In the next experiment, the bananas were still further from the cage, out of reach of hand or stick, but inside the cage were not one but two short sticks, which could be slotted together to make a longer one. In what was perhaps the most famous experiment of the series, Köhler's ape, after trying and failing to reach the bananas with the short stick, finally 'hit upon' the possibility of slotting them together and immediately succeeded in dragging the bananas to himself. In a primitive sort of way Köhler's ape, by making a 'creative synthesis' in order to bring about a desired result and solve a particular problem, was performing that class of activity which one might regard as characteristically human and particularly 'scientific'.

1.2.5 Generalization

Another type of human behaviour which could be regarded as 'scientific' (in the sense of definitions (i) and (ii) on p. 11) is also shown by other animals. This is the capacity to generalize. If a chimpanzee is set a problem in which he is presented with a row of cubes, under the third of which there is a peanut, he will eventually learn to look without error under the third cube for the peanut. If the cubes are replaced by some other shape, say pyramids, the animal will once more learn that the peanut is to be found under the third, and by the time the pyramids have been replaced by yet other shapes of objects the animals will learn to generalize, that whatever the object, the peanut is always under the third in line. This process of *generalization* is very much what one part of science is about (it is an example of *induction*, see p. 44).

What we have tried to show so far is that, although the activities described as 'science' are unique to man, there are in the animal world a whole series of activities in which the primitive analogies to modern scientific activity can be recognized. Human science and technology, which have produced relativity theory and the understanding of the genetic code, penicillin and hydrogen bombs, may seem far removed from such pre-human examples, yet they are clearly related.

Of this particular evolutionary line, of increasing brain power, increasing power of analysis, exploration, learning and manipulation, man has so far been the outstanding end-product. The arrival of man on the world scene meant the coming of an animal which in at least three major respects was immensely more accomplished than any predecessor.

Suggest which these might be.

Ability to think.
Ability to speak.
Ability to make tools.

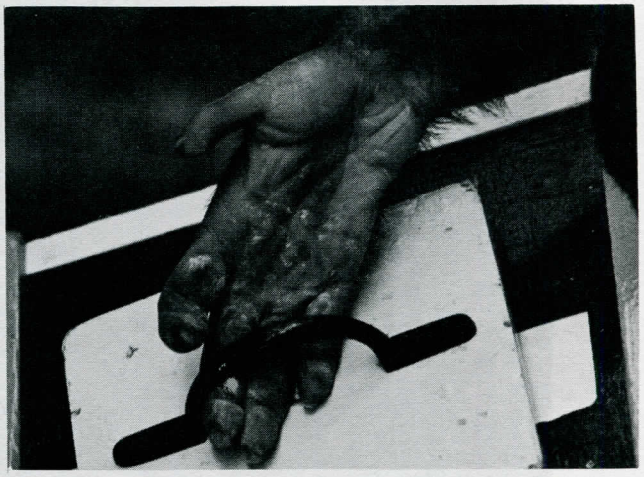
1.2.6 Tools

The last of these accomplishments is the easiest to understand. Human hands are different even from those of man's nearest evolutionary neighbours, particularly in the position and manoeuverability of the thumb. Compare, for example, Figures 4a and b. Human thumbs can be swivelled into a variety of positions relative to the fingers impossible even for chimpanzees. This makes holding and adjusting things possible, makes fine control of delicate objects feasible, and in its turn leads to the possibility of manufacturing hand-extensions, i.e. tools far removed in complexity and ingenuity from the sticks, stones and leaves used by several kinds of ape. In addition, man, unlike the apes, can stand upright on two legs without the need to use his hands to help with balancing. This frees his hands for other functions. The gap between the contemporary tape- and computer-controlled machine tool, and the hand-chipped flint axe is less,

perhaps, than that between man's flints and anything that had gone before. The development of tools, in the sense of extensions to man's hand, and instruments in the sense of extension to his senses, thus characterize man's emergence as *Homo sapiens*. However, the use, and to an even greater degree the development, of tools is not really possible for a man in isolation, but depends upon the social nature of humans, the fact that they live in communities.

Figure 4a

Figure 4b



1.2.7 Speech

The second major feature to arrive with man is the power of speech. Not even the most advanced of apes has this capacity. Certainly many animals can signal quite complex messages to each other. Bees can signal not only the presence but also the direction and approximate distance of nectar. Birds can inform one another whether they are paired or single, whether one bird is a trespasser on another's territory, and so forth. Distress, danger and sexual willingness can be signalled by many animals. Dolphins can apparently communicate by sound signals. But only man has the ability to represent the steps of an argument or to relate a sequence of events by means of an ordered sequence of spoken or written signs.

It is often said that parrots, budgerigars and some other birds can speak. Is this true?

No—such birds can imitate sounds that they can hear, and can often do this with striking accuracy, but there is no suggestion that they can use the sound to signal meaning or ideas, or to communicate the content symbolized by the sounds.

Nor can any animal draw object-like pictures. Although some apes will paint rather unstructured patterns, they are never recognizable pictures of objects. The earliest of these are the cave paintings of man (compare Figs. 5a, b, and c for instance, pp. 20 f.).

Much effort has been expended in trying to teach apes to talk, and it is not altogether clear why they cannot. In 1969 an experiment was described in which a chimpanzee was reared from birth in the home of an experimental psychologist who claimed to have been able to teach the animal a sign language of a 'few dozen' words, similar to those used in communicating with the deaf-and-dumb. But interesting as such experiments are, it is unlikely that the chimpanzee will ever advance beyond the vocabulary of a human child of one or two. There are some suggestions that dolphins can communicate both with each other and with humans, but once again, the extent of this communication is very limited. Communication by speech is unique to the human species.

Speech made possible the transfer of experience between individuals, and

Figure 5a *Cave painting.*

Figure 5b *Picture by a six-year-old child.*

Figure 5c *Picture by a chimpanzee.*

the subsequent development of written communication systems extended this transference to individuals who might be distant from one another in space or in time. Thus, instead of having laboriously to assemble a 'map of the world' from scratch with perhaps a few innate and stereotyped rules, like those of recognition of 'mother' and 'danger' we described earlier for birds, each generation of man has been able to build upon the experience of those which came before. A discovery once made could be recorded, elaborated upon and improved, a process itself speeded up by the development first of writing, then printing, and more recently of film, tape and computer.

Such stores of knowledge, experience and new data represent extensions of man's brain capacity, that is, of his ability to think, which we noted above as the most important new feature which distinguishes man from his evolutionary forbears. Even without them he could carry in his head vastly more information than could other animals as a result of the more complex structure of his brain. With the addition of these new extensions, as we pointed out at the very beginning of this Unit, his communication and handling of data can be speeded by factors of a million or more. It is therefore by cashing in on and exploiting the three evolutionary advantages with which he started (those of brain, speech and tools), that man owes his present evolutionary success. Science and technology represent the systematic exploitation of these advantages.

1.3 The birth of science

With the appearance of the distinctly human powers of thought, speech and tool-making, the evolutionary development of man took a new turn: from *biological* to *social*. Although other species of animals have sensory equipment and skills that are quite impressive, and which often exceed their counterparts in humans, all the evidence indicates that the range of abilities for each species is fixed. The experience of one generation cannot modify substantially the behaviour of the next. Certain animals have been capable of being 'tamed' and of interacting with their human hosts in a very sophisticated way, but a twentieth-century cat probably has no greater range of skills than one that the ancient Egyptians worshipped. By contrast, within the limits of its genetic constitution, the human species has been adding to its stock of knowledge and skills at a constantly increasing rate.

Some of the crucial inventions that brought human life further and further from that of the animals are well known: the use of fire; the control of domestic animals by, for instance, the invention of the bridle and stirrup; the invention of the wheel; the manufacture of cutting tools and weapons; pottery; agriculture; smelting and alloying of metal. There were doubtless many others, from earliest times, of which no material trace remains. As new inventions were *achieved* and *retained* through successive generations, social life itself was transformed. Hunters preying on animals became herdsmen and stock-breeders; people living in a more settled way, gathering edible fruits and grasses, became farmers; and the possibilities inherent in a settled society with a social division of labour were realized in the great urban civilizations of Egypt, Mesopotamia, India and China.

In the achievement and transmission of these basic techniques, there was an element of 'science' present. For the attempts at the *understanding, prediction and control* of nature went hand in hand. Moreover, each of these tasks involved, in its own way, a direct contact with the uncontrolled world of nature. The regularities of the natural world, the cycle of the seasons and the Moon, the rising and setting of the Sun, the revolutions of the stars in the heavens, were important for huntsman and farmer alike. Irregular phenomena, usually natural disasters such as floods and earthquakes, storms and drought ('acts of God', as they are still called in a Householder's Comprehensive Insurance Policy) but also strange occurrences like eclipses and rainbows, called for an understanding of their causes, and, if possible, prediction. And since men then believed that all these events (the regular and the irregular) were caused by personal agents (gods) who needed to be honoured and pleased, there was no sharp distinction between religion, magic, technology, and 'science'.

Quite late in human social evolution there developed what has become known as 'pure science': the quest for knowledge for its own sake. This has been called the 'Greek miracle' of the fifth century B.C. In that period were established the ideas of mathematics as a system of *deductive knowledge* rather than a technique for measuring and reckoning; of philosophy as an enquiry into the nature of human knowledge; and of the explanation of natural phenomena independently of myths and magic. Certain sciences were found to be capable of systematic development as coherent bodies of knowledge, and in the following centuries grew to be independent of their roots in ancient craft practice. Perhaps the first was mathematics. There followed the emergence of astronomy as distinct from astrology; of anatomy and physiology from practical medicine; and botany from pharmacy and agriculture. But these developments were long and uneven. The ideal of 'pure science' did not get firmly established

until nearly the nineteenth century; and even today there are more people in the United States—and perhaps in Britain—earning a living from astrology than from astronomy!

Even after the birth of 'science' as modern men understand it, technology and the practical crafts were at first only slowly and patchily affected by science, but continued to evolve in response to their own needs and possibilities. For knowledge of the *causes* of natural phenomena was until recent times too limited (and usually too unreliable!) to guide the process of invention of new means of controlling their *effects*. Even in China, where there was a long tradition of sophisticated technology and of enquiry into the natural world, scientific research as we understand it never got under way. And in other societies such as those of Ceylon and Mesopotamia quite elaborate systems of waterworks could be invented and operated in the absence of 'science'. Through many centuries, and in all civilizations, there tended to be a division of labour: the causes of natural events were studied from books and was an occupation for the educated scholar; while control of effects was left to the artisan, who got his hands dirty, and was considered to need little, if any, literacy. This social division was clearly marked in medicine in Europe: the university-trained physician would diagnose and prescribe on the basis of the patient's horoscope and by examination of the colour of his urine (this practice is typified by the Doctor of Physick in Chaucer's *Canterbury Tales*), while the barber-surgeon performed particular operations learned as craft secrets from a master.

1.3.1 The Renaissance

The unification of 'science' and 'arts', 'craft' and 'technology' began to take place in Europe during the Renaissance in the fifteenth and sixteenth centuries. An increasing range of crafts received the attention of learned men (navigation, warfare, mining, surgery) and scholars gradually came to use and value craft experience, to develop the idea of making *controlled experiments*. It was at this time, too, that the sort of 'science' with which we are now familiar began to emerge out of this contact and mutual enrichment between the work of scholars, craftsmen and artists. By the beginning of the seventeenth century, work which has become a permanent part of our accepted scientific knowledge, such as, in England, Gilbert's discovery of terrestrial magnetism and Harvey's discovery of the circulation of the blood and, in Italy, Galileo's discovery of the laws governing the motion of falling bodies, was appearing.

Very soon there was established an approach to enquiry in the natural world which we can see as 'science', in many of the senses defined on p. 11.

Do you recall the definitions of 'science' we adopted there?

If not, check back!

In this new sort of science, there was a much closer relationship with technology: from industrial (or military) practice would come ideas for problems and techniques and instruments; and when solutions were achieved, every attempt would be made to turn them to profitable use.* By the middle of the seventeenth century, the methods of science as we

* *We have not time here to trace the history of subsequent scientific developments. For further reading on this point, see Bernal,* Science in History, *Chapter 7, and Rose and Rose, Chapters 1–6.*

now know them were beginning to be formulated explicitly, although its social organization was to undergo further changes. From then onwards, the accepted approach to the study of nature was based on controlled experience (systematic observation or experiment), in conjunction with careful argument, for the deriving of conclusions about the behaviour of the natural world; and such results of research were submitted to the community of scientists for their assessment and use. In its results and in its methods, modern science represents the highest point yet reached in the development of man's cognitive faculties. Although science will doubtless evolve further in directions we cannot predict, and to the outsider does not seem to involve the 'affective' domain of emotion and imagination, it defines and moulds our civilization and ourselves.

1.4 Scientific knowledge and craft knowledge; an experiment

Even in this modern technological age, we find ourselves using, in our ordinary lives, craft skills that are learned by imitation and practice, and we do this even for the operation of equipment which depends on science for its design and production. One example is driving a car; one needs to know almost nothing about the principles of an internal combustion engine to do this successfully in ordinary use. Another example is doing household electrical jobs: but here the handyman must know something about electricity if he is not to make disastrous mistakes.

Before going on to consider the differences between craft knowledge and scientific knowledge, we want you to try an experiment, for there is no better way of learning about science than to be involved in solving a specific problem. In particular we want you to repeat one of the most famous experiments in the history of science. In it, you will see the relationships between craft and scientific skills demonstrated. The experiment was done by the great Italian scientist Galileo Galilei, who lived from 1564 to 1642 and whose work led to the discovery of the laws governing the movement of bodies and to the demonstration that the Earth moved round the Sun. Galileo invented many instruments; he pioneered the telescope, for example, and his work led him into conflict with the Catholic Church of his time, and to a famous and epoch-making dispute between scientific truth and established authority whose echoes still ring today. He is often erroneously credited with having refuted Aristotle's views on the laws of motion by dropping two weights from the leaning Tower of Pisa. The reality may have been less dramatic, but was certainly more interesting. When Galileo began his experiments, there was almost complete confusion in people's minds about such concepts as force, motion, velocity and acceleration. Galileo helped clarify these concepts (it was left to Isaac Newton finally to produce the mathematical definitions of gravity and of force and motion that were to serve as fixed points of scientific reference for over 200 years, until Einstein's formulation of relativity theory at the beginning of the twentieth century). But Galileo did more than just design ingenious experiments. More almost than any other man he developed experimental method and procedure, ushering in a new era in human thought—that of systematic science. In following his experiment, you will yourself see and feel this procedure in operation. You will also encounter another aspect of science. For, as you will discover, even great scientists can make mistakes!

1.4.1 Galileo's inclined-plane experiment

In this experiment, made in the first years of the seventeenth century, Galileo established experimentally that the distance travelled by a descending body increases in proportion to the square of the time that has elapsed. Most schoolboys who have studied physics have 'confirmed' Galileo's results by doing the experiment themselves. You may well wonder why this new science course of the Open University should start with such an old, standard and 'hackneyed' experiment. We have chosen it deliberately; for in many ways this experiment can illustrate more about scientific research and scientific method than the most complicated and exciting modern apparatus.

First, this is not the version of 'Galileo's experiment' demonstrated in schools. The inclined-plane experiments that are designed for schools' use are constructed to be foolproof, or 'student-proof'. They enable the

unskilled student to get accurate results, and relieve him of the necessity of grappling with the conceptual and practical problems that Galileo had. They are descended from Galileo's own experiments, but until now, to our knowledge, no science course has attempted to teach the experiment as nearly as possible to the way in which Galileo did it himself.

Even when it is done in Galileo's own way, the experiment is simple. There is not much equipment involved, and the measurements are straightforward. Hence you should have little difficulty in seeing what *ought* to be happening; and we will have more time available for discussing difficulties in the experiment itself and the problems of significance.

In the discussion you will learn some things about the experiment that may be surprising. For the experiment has, for many years, been considered as something very simple and obvious; so simple that it needed no 'hypotheses' for its design, and no 'theory' for its interpretation. After all, if one wants to know how bodies increase their speed as they fall, what is more natural than to let them drop and see what happens? Galileo's only merit was that he was one of the first people to approach the problem in a scientific way, rather than trying to find the answer in the then 2000-year-old writings of Aristotle. And this famous example has given the impression that the first steps in any science are simple experiments, free from theory, out of which quantitative laws are derived.

We want you to try the experiment first before talking about it too much; but you can immediately notice one feature of it that makes it not quite the obvious solution to the problem. Galileo was concerned with the behaviour of falling bodies and with the relationship between the time of descent and the speed at the end of the descent. This was a problem of great philosophical and scientific interest in his time; and later his law of *falling* bodies was generalized by Isaac Newton to form the theory of universal gravitation, which in turn became the foundation of celestial mechanics, and of astronautics. But if you look at the experiment, you will see that the bodies are not made to *fall*, but to *roll*. This, as you will see later, is a crucial difference. It is clear today and it was clear to Galileo that there are important differences. And so this experiment, by itself, does not give us a certain answer to the problem of the behaviour of falling bodies. Hence Galileo needed other experimental results, which we will explore later.

1.4.2 How Galileo described the first part of his experiment

Let us now read how Galileo described the experiment:

> A piece of wooden moulding or scantling, about 12 cubits long [*Note*: 1 Florentine cubit was about 0.6 metres, so the total length of the piece was about 7 metres], half a cubit wide, and three finger-breadths thick, was taken; on its edge was cut a channel a little more than one finger in breadth; having made this groove very straight, smooth, and polished, and having lined it with parchment, also as smooth and polished as possible, we rolled along it a hard, smooth, and very round bronze ball. Having placed this board in a sloping position, by lifting one end some one or two cubits above the other, we rolled the ball, as I was just saying, along the channel, noting, in a manner presently to be described, the time required to make the descent. We repeated this experiment more than once in order to measure the time with an accuracy such that the deviation between two observations never exceeded one-tenth of a pulse beat.

(*Note:* Galileo here reports sound experimental practice, that of making sure that the equipment is working consistently, before proceeding to try variations on the experiment. His report of agreement to within one-tenth of a pulse beat sounds exaggerated, and may well have been so. His time-measuring device was probably capable of such precision, as we shall see below; and once he got 'in tune' with the apparatus, he may well have achieved that sort of accuracy.)

> Having performed this operation and assured ourselves of its reliability, we now rolled the ball only one-quarter the length of the channel; and having measured the time of its descent, we found it precisely one-half of the former. Next we tried other distances, comparing the time for the whole length with that for the half, or with that for two-thirds, or three-fourths, or indeed for any other fraction; in such experiments, repeated a full hundred times, we always found that the spaces traversed were to each other as the squares of the times, and this was true for all inclinations of the plane, i.e. the channel, along which we rolled the ball.

(*Note:* He repeated readings in order to get reliable results. This is also an example of good experimental technique.)

> We also observed that the times of descent, for various inclinations of the plane, bore to one another precisely that ratio which, as we shall see later, the Author had predicted and demonstrated for them.

(*Note:* This passage describes quite a different experiment, which, as we shall see, was quite crucial to Galileo.)

> For the measurement of time, we employed a large vessel of water placed in an elevated position; to the bottom of this vessel was soldered a pipe of small diameter giving a thin jet of water, which we collected in a small glass during the time of each descent, whether for the whole length of the channel or for a part of its length; the water thus collected was weighed, after each descent, on a very accurate balance; the differences and ratios of the times were measured, and this with such accuracy that although the operation was repeated many, many times, there was no appreciable discrepancy in the results.

(*Note:* This bit of experimental technique is a good example of Galileo's genius at precision instrumentation. Although there were clocks in his time, they were not designed for the accurate measurement of small intervals of time; and even the pendulum, whose use was developed by Galileo, could not give sufficient precision. Knowing that weight could be measured to a greater precision than any other physical quantity, Galileo devised this method of 'weighing the time'.)

On the question of accuracy, we can calculate that the rolling ball would take about five seconds for a full descent; and 'one-tenth of a pulse beat' would be about a hundredth of that. If Galileo's 'pipe of small diameter' discharged at 25 $cm^3\ s^{-1}$*, a full run would produce water weighing some 125 g. This would need to be weighed accurately to, say, 0.5 g; and the scales for this would be in use among goldsmiths (whose trade Galileo knew very well: his most famous argument was published as *Il saggiatore* —'he who weighs the evidence'). Hence the largest source of inaccuracy would be the timing of the start and finish of the descent. So any improvement of Galileo's apparatus was not likely to improve the accuracy of the experiment.

1.4.3 The first experiment: what you have to do

You cannot reproduce Galileo's experiment exactly; it would be difficult to set up a track of the same size, and measuring many time intervals by weighing is time-consuming and tedious. But we hope that you have a table more than a metre long, which you can raise at one end. You have been provided with a stopwatch and a ball. The question of how much to raise the end of the table to get a suitable tilt is an example of experimental design.

What are the disadvantages of making it too steep, or too flat?

Too steep: the ball rolls too quickly and is difficult to time, also a strain is exerted on the table legs.
Too flat: the ball's descent may be affected by roughness in the surface of the table.

**that is, cubic centimetres per second*

Arrange for the gradient to be such that it takes about 2 to 4 seconds for a ball to roll the full distance of about 1 metre. For this part of the experiment there is no need to measure the gradient.

The ball may not roll down the table in a straight line. Does this matter?

You could answer this question experimentally or theoretically. The experiment approach would be to observe whether the ball does "wander" off a straight path down the slope—for instance because of a sideways tilt of the table, or because the table top is not flat. If it does wander off, does this affect the measured time it takes to cross the "finishing line" by an amount that is significant compared with the time and with errors in measuring that time?
The theoretical approach might be to calculate the extra time it would take if it deviated by a given amount from the proper path. For example, if the ball deviates by as much as 0.3 m to either side of a straight line 1 m long, the error in the distance it rolls will be less than 10 per cent, and in the time, less than 5 per cent. How much is 5 per cent of 2 seconds? Can you measure this with your stopwatch? Do you get the same time from one measurement to the next, even if the ball rolls down more or less the same path each time?

We want you to time how long it takes the ball to cover the full distance. For timing the runs, you will want to use the stopwatch. You might try doing the timing with a sweep-hand of a wrist watch; you will see that (unless you are very skilled) this instrument is not sufficiently accurate for the job. But even using this modern equipment does not take all the fun out of the experiment. For you will need to train yourself to get an accurate timing at the beginning and end of the descent. In many modern versions of the experiment, there are devices for automatically starting and stopping the clock; these make it easier, for their inherent inaccuracy is far less than that of other parts of the experiment. But, with a manually operated timing device, you are at the limits of accuracy of your experiment, just as Galileo was. You can practise starting the watch just as you release the ball, and stopping it just as the ball strikes something at the terminus that, preferably, gives a noticeable sound. You should also give some thought to the best way of releasing the ball—having tried it a few times by holding and releasing it from your fingers you may consider this not good enough. (What if your fingers vary in their degree of 'stickiness' during the experiment?) Also you might consider wiping the table top with a duster before you start. Only when you have these various factors under control and repeated readings are giving reasonably reliable results should you start recording the measurement in earnest. Note that this is just what Galileo did. Having taken the average of the readings and recorded the distance, you should time the runs over a series of shorter distances. Record your readings in a notebook in the form of a table, as shown in the example. (Don't worry about the final column of the table—that will be added later.)

Distance of run (cm)	Time of run (seconds)						The square of the average time (s^2)
	Individual runs					Average time	
	1st	2nd	3rd	4th	5th		
90	3.0	2.9	2.9	3.0	2.8	2.9	8.4
75	2.4	2.6	2.6	2.6	2.3	2.5	6.25
60	2.3	2.4	2.2	2.3	2.3	2.3	5.3
45	2.0	1.8	1.8	2.2	2.0	1.95	3.7

Carry out the experiment as described.

Having recorded your readings we want you now to pause and take stock of the situation.

Note first we have already given you a number of general tips—perform a preliminary experiment to finalize the design, take an average of repeated readings, record results in a notebook, show your readings in tabular form. We do not want to have to repeat these tips every time you come to do an experiment. We have therefore produced a special book entitled *The Handling of Experimental Data*, which contains, among other things, this general type of advice. We shall refer to this book as *HED*. If you

have had little previous experience of practical work you will probably find *HED* rather difficult to digest in one session—we shall therefore direct you to certain selected sections of it as and when the occasion arises.

Having collected your raw data, i.e. the distances and their corresponding times, we want you now to verify the result found by Galileo. This involves calculating the squares of the average times (i.e. filling in the final column to the table) and plotting a graph of distance against (time)^2. If the distance is indeed proportional to the square of the time, this graph should yield a straight line passing through the origin.

HED:
(i) In order to learn what this book contains and why we consider it important for you to do practical work, read the introduction.
(ii) If you had any difficulty in knowing how and why you should take the *average* of the set of timings, read *HED* 2.1.
(iii) Read *HED* 2.12 on the need to keep a notebook and 2.13 which gives hints on performing calculations.

Read *HED* 2.7, 2.8, 2.9, and 2.10 which explain proportional relations, and why graphs are considered important, and gives hints on how to plot them and extract information from them. (In *HED* 2.9 reference is made to the calculation of the error on the slope of a graph. You need not worry now about this point as this requires a knowledge of earlier sections.)

Now plot your graph and verify Galileo's distance-time relation, as quoted on p. 27.

1.4.4 Introduction to the second part of Galileo's experiment

By this point you know that you should have obtained sets of readings which fit the relation: distance proportional to the square of the time. But this applies for a ball ROLLING down an incline; *how can we know that it also applies to the same ball falling freely*? Timing how long it takes the ball to drop freely through various fractions of a metre would be tricky—the time intervals would be uncomfortably short. Galileo recognized the difficulties of performing such a direct experiment. So he thought of approaching the problem indirectly by using his inclined plane. He repeated the experiment using steeper inclines (as he says in his description). He found that the relation held at any inclination (*if you wish* you can repeat the first experiment at different inclines to satisfy yourself that this is so). He then argued that the relation would hold in particular for the limiting case of a vertical incline. By measuring the distance the ball rolls in a given time down planes of various inclinations, he hoped to extrapolate to get a distance rolled down a vertical plane. This he was to equate to the distance a ball would travel in free fall in the given time. The question was—how to extrapolate to a limiting case? Look at Figure 6.

AB is a vertical diameter of a circle, CB is a chord, and C′B is another chord, longer and steeper than CB.

Galileo guessed that the time a body would take to roll from C to B would be the same as the time from C′ to B. If this hunch (and it was nothing more) could be verified experimentally for several different chords of the circle, like CB, C′B, C″B, etc., it would seem reasonable to suppose it might hold good all the way up to the vertical chord AB. This is called the 'circle-chord theorem'. Galileo argued that if the distance-time relation is found to hold for different lengths *and* different slopes, even very steep, 'almost vertical' slopes, then it must hold also for the vertical. In this way he solved the problem of extrapolating from the measurable case of descent along a relatively shallow incline to the non-measurable case of free fall.

You may be wondering how Galileo got his hunch. The history is obscure, but it seems that his experiments in mechanics started, not with the inclined plane, but with the *pendulum*. Very early in his scientific career he did experiments with the pendulum. He found the time for completing a back-and-forth swing to be independent of the amplitude of the swing, i.e. how far it swings. If you take the pendulum bob as a point falling and rising along a circular track, then the time of descent along all *arcs* of a circle to the lowest point are the same provided the arcs are small. Roughly speaking, the longer the arc, the more steeply inclined it is at the extremity, the more quickly the body starts its descent, and so makes up for the extra distance to be travelled to the bottom. The circle-chord theorem probably occurred to Galileo by analogy, when he had shifted

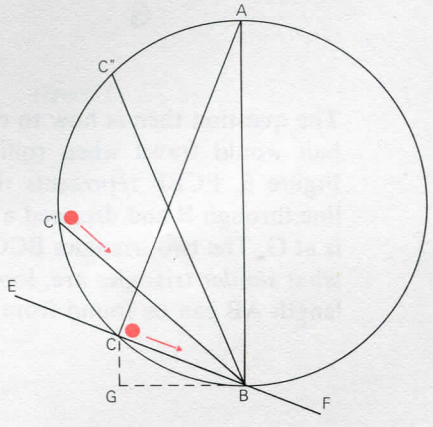


Figure 6

his attention to the problem of free fall, and used the straight incline for experiment. As it happens, while the circle-chord theorem is true, the circle-arc theorem is only approximately true for small arcs.

The next experiment we want you to do is to *verify the 'circle-chord theorem'*, i.e. for a number of inclines measure the distances travelled in a fixed time of one second and check that they correspond to chords of a circle (like CB, C′B, C″B). From the calculated diameter of this circle you can then extrapolate to the vertical case, i.e. calculate the length of the chord AB which, as Galileo assumed, should be the distance the ball will travel in free fall in one second. Thus, on Galileo's reasoning, *you will have arrived at a measure of the acceleration due to gravity.*

1.4.5 The second experiment: what you have to do

Measure the time for the ball to roll down various lengths of track, keeping the slope of the table fixed. By trial and error find the distance that corresponds to a one second time of descent. This distance is CB in Figure 6.

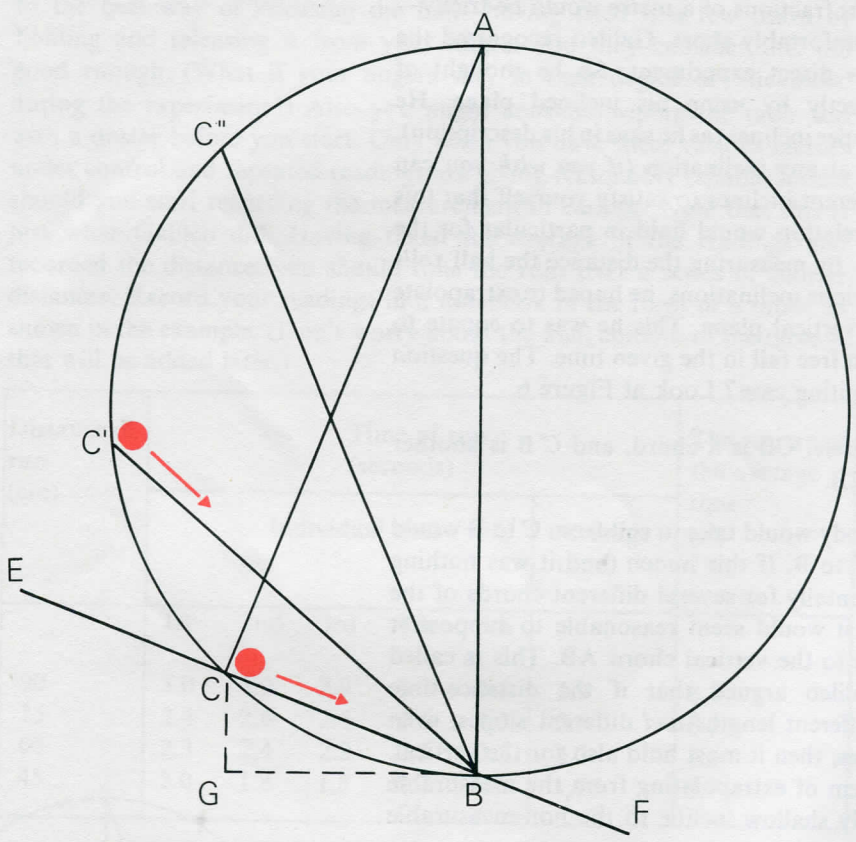


Figure 6

The question then is how to calculate the distance AB, i.e. the distance a ball would travel when rolling down a vertical slope. Look again at Figure 6. ECBF represents the table top. We have drawn a horizontal line through B and dropped a perpendicular to it from C; the intersection is at G. The two triangles BCG and ABC are similar. (If you do not know what *similar* triangles are, look at *MAFS*, section 2.E.) Consequently the length AB can be found from the relation

$$\frac{AB}{CB} = \frac{CB}{CG} \quad \text{.....(1)}$$

(Corresponding sides of similar triangles have the same ratio.)

have had little previous experience of practical work you will probably find *HED* rather difficult to digest in one session—we shall therefore direct you to certain selected sections of it as and when the occasion arises.

Having collected your raw data, i.e. the distances and their corresponding times, we want you now to verify the result found by Galileo. This involves calculating the squares of the average times (i.e. filling in the final column to the table) and plotting a graph of distance against (time)^2. If the distance is indeed proportional to the square of the time, this graph should yield a straight line passing through the origin.

HED:
(i) In order to learn what this book contains and why we consider it important for you to do practical work, read the introduction.
(ii) If you had any difficulty in knowing how and why you should take the *average* of the set of timings, read *HED* 2.1.
(iii) Read *HED* 2.12 on the need to keep a notebook and 2.13 which gives hints on performing calculations.

Read *HED* 2.7, 2.8, 2.9, and 2.10 which explain proportional relations, and why graphs are considered important, and gives hints on how to plot them and extract information from them. (In *HED* 2.9 reference is made to the calculation of the error on the slope of a graph. You need not worry now about this point as this requires a knowledge of earlier sections.)

Now plot your graph and verify Galileo's distance-time relation, as quoted on p. 27.

1.4.4 Introduction to the second part of Galileo's experiment

By this point you know that you should have obtained sets of readings which fit the relation: distance proportional to the square of the time. But this applies for a ball ROLLING down an incline; *how can we know that it also applies to the same ball falling freely*? Timing how long it takes the ball to drop freely through various fractions of a metre would be tricky—the time intervals would be uncomfortably short. Galileo recognized the difficulties of performing such a direct experiment. So he thought of approaching the problem indirectly by using his inclined plane. He repeated the experiment using steeper inclines (as he says in his description). He found that the relation held at any inclination (*if you wish* you can repeat the first experiment at different inclines to satisfy yourself that this is so). He then argued that the relation would hold in particular for the limiting case of a vertical incline. By measuring the distance the ball rolls in a given time down planes of various inclinations, he hoped to extrapolate to get a distance rolled down a vertical plane. This he was to equate to the distance a ball would travel in free fall in the given time. The question was—how to extrapolate to a limiting case? Look at Figure 6.

AB is a vertical diameter of a circle, CB is a chord, and C′B is another chord, longer and steeper than CB.

Galileo guessed that the time a body would take to roll from C to B would be the same as the time from C′ to B. If this hunch (and it was nothing more) could be verified experimentally for several different chords of the circle, like CB, C′B, C″B, etc., it would seem reasonable to suppose it might hold good all the way up to the vertical chord AB. This is called the 'circle-chord theorem'. Galileo argued that if the distance-time relation is found to hold for different lengths *and* different slopes, even very steep, 'almost vertical' slopes, then it must hold also for the vertical. In this way he solved the problem of extrapolating from the measurable case of descent along a relatively shallow incline to the non-measurable case of free fall.

You may be wondering how Galileo got his hunch. The history is obscure, but it seems that his experiments in mechanics started, not with the inclined plane, but with the *pendulum*. Very early in his scientific career he did experiments with the pendulum. He found the time for completing a back-and-forth swing to be independent of the amplitude of the swing, i.e. how far it swings. If you take the pendulum bob as a point falling and rising along a circular track, then the time of descent along all *arcs* of a circle to the lowest point are the same provided the arcs are small. Roughly speaking, the longer the arc, the more steeply inclined it is at the extremity, the more quickly the body starts its descent, and so makes up for the extra distance to be travelled to the bottom. The circle-chord theorem probably occurred to Galileo by analogy, when he had shifted

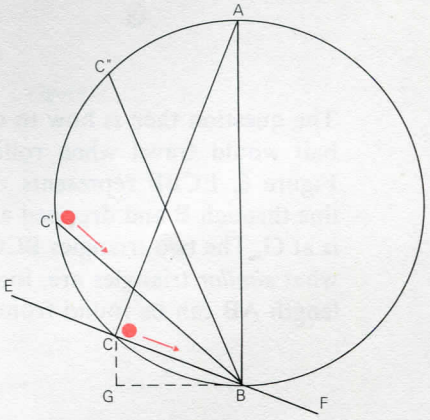


Figure 6

We already know CB. It is the length for a one second descent down this particular slope. If we can somehow measure CG then we have AB.

Can you think of a way of measuring the ratio CB/CG without using a spirit level and a plumb line?

One way of doing it is suggested by Figure 7. Here we have drawn in the table, legs and all.

Then, provided the table top is flat and its legs are of equal length, the triangles CGB and HKJ are similar so,

$$\frac{CB}{CG} = \frac{HJ}{HK}$$

You can measure HJ and HK directly (you will see that you do not really need a plumb line for this, for the error caused by slight deviations from the vertical is small). So now you have

$$AB = CB \times \frac{HJ}{HK} \quad \text{.....(2)}$$

and everything on the righthand side of the equation has been measured. You can now estimate the diameter of the 'circle of one second fall time'.

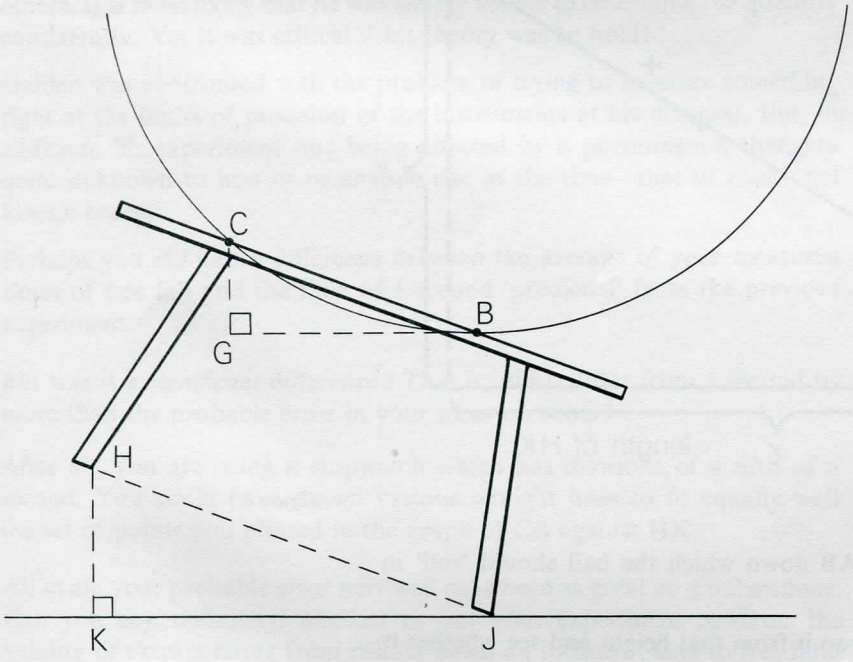


Figure 7

Next, change the tilt of the table and find the value of CB for the new inclination, i.e. the new distance travelled in one second (the distance C′B in Figure 6). Also measure the corresponding value of HK. If the 'circle-chord theorem' is correct then the value of AB obtained from equation 2 with these values of CB and HK should agree with the value obtained with the first inclination of the table, i.e. CB and C′B should be chords of the same circle.

Repeat the experiment for a total of at least five different inclinations.

Substituting the readings in equation (2) is certainly one way of verifying the 'circle-chord theorem'. In *HED* 2.8 you were advised to display results in the form of a graph whenever possible. How could you do so with your present results?

In order to see how you can display your results graphically we rearrange equation (2) so:

$$\frac{AB}{HJ} \times HK = CB \quad \ldots\ldots\ldots\

Do not read on until you have done this experiment and answered these questions:

Did you get a time of 1 second? Did you expect to?

You should, in theory, have measured a free fall time of about 0.8 seconds.

The difference, 0.2 seconds, is about the same as the probable error in your measurements.

You would need a more precise timing system to check the validity of this hypothesis.

Perhaps now you will appreciate Galileo's dilemma. His timing system was, if anything, inferior to yours. He could measure *relative* times quite accurately with his water clock, but to measure absolute times, times in seconds, he had to compare some sort of continuously repeating process with the length of the day or with some other astronomical measure of time. The best available device of this kind was the pendulum, but the 'pendulum clocks' of Galileo's time were exceedingly primitive and inaccurate; indeed it was not until 1657, fifteen years after the death of Galileo, that the Dutch scientist Huygens invented an effective pendulum clock.

Galileo was well aware of the importance of defining physical constants and measuring them; yet for this most important one, the height of one-second free fall, his estimates are given casually, and are not consistent. In some places, they seem to indicate a value a bit higher than the correct one; in others, they indicate the value obtained by extrapolation from the inclined-plane; but he never committed himself in this case as he did in others. It is most likely that he was simply *unable* to determine the quantity consistently. Yet it was critical if his theory was to hold!

Galileo was confronted with the problem of trying to measure something right at the limits of precision of the instruments at his disposal. But, in addition, his experiment was being affected by a phenomenon that was quite unknown to him or to anyone else at the time—that of rotational kinetic energy.

Perhaps you *did* find a difference between the average of your measured times of free fall and the time of 1 second 'predicted' from the previous experiment.

But was it a *significant* difference? That is, did it differ from 1 second by more than the probable error in your measurements?

After all, you are using a stopwatch which has divisions of a fifth of a second. You could have drawn various straight lines to fit equally well the set of points you plotted in the graph of CB against HK.

All in all, your probable error may well have been as great as ± 0.2 seconds. Can you say, definitely, whether or not your experiment confirms the validity of extrapolating from rolling down an inclined plane to free fall?

Probably you cannot. If you care to look up Appendix 2 (Black) you will see that the effect of *rotation* on a rolling ball is to use up some of the energy that would otherwise be available to give it linear kinetic energy down the inclined plane. The result is that it will take about 20 per cent longer to *roll* a given distance than to *slide* the same distance on a friction-less surface. As a result, the time the ball should take to *fall* the height AB should be about 20 per cent less than the time you would predict by the extrapolation you did.

The importance of making due allowance for the energy going into rotation can be demonstrated in a very marked fashion by considering a pair of

Summary of the Galileo Experiments

Apparatus required

A table (more than 1 metre long)
Stopwatch
Ball
Ruler or tape-measure

Experiment 1

Tilt the table so the ball takes 2 to 4 seconds to roll 1 metre.
Time the run over about half-a-dozen distances taking at least five readings for each distance.
Average the time for each run and square it.
Plot distance against (time).[2] Verify that it is a straight line passing through the origin.

Experiment 2

Vary the slope of the table, and for each slope measure the distance covered by the ball in one second, and the distance HK of the inside of the raised table leg above the level of the floor (Fig. 7, p. 31).
Plot a graph of the distance the ball rolls in one second against HK.
Estimate the slope of the straight line obtained. Measure the distance HJ between the feet of the legs of the table, and hence determine the height of AB (Fig. 6, p. 30) through which the ball would roll down a vertical plane in one second.

Test of the Experiment

From a height equal to this estimated value of AB drop the ball several times and see whether it does indeed take one second to strike the floor.

Final Report

This should be written for submission to your tutor.

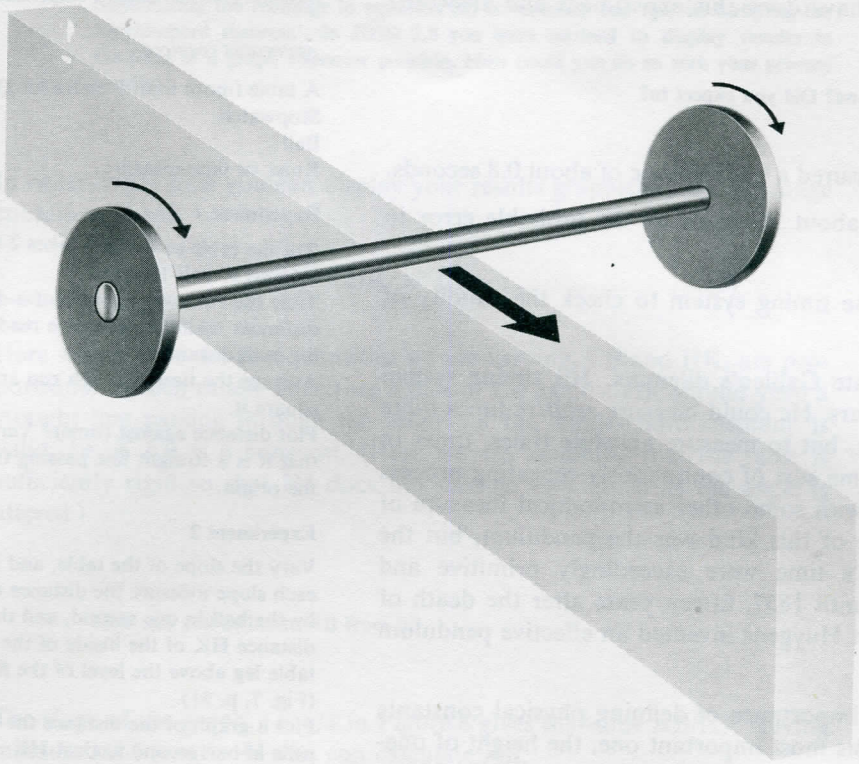


Figure 9

heavy wheels on an axle, rolling down an incline on its *axle* as in Figure 9.

As you can probably well imagine it would take much longer for this object to *roll* to the bottom than to *slide* the same distance.

Another way of seeing what happens is to look at Figure 10, where a ball is shown rolling down an almost vertical incline.

In the case shown, the almost vertical frictional force is necessary in order to impart rotation to the ball—but the same force also partly cancels out the gravitational force; in so doing it reduces the rate of descent compared to the case where the ball falls freely with no increase in its rate of rotation.

If you have done the inclined-plane experiments at school, you may have noticed that the weight is usually carried in a little cart, with light wheels to avoid this problem, and in advanced mechanics, the kinetic energy of rotation of the wheels is taken into account when the calculation is made. In this way, the modern student is led past the pitfalls that Galileo encountered. But how much more interesting it is if, at least once, we can see the mixture of truth and error which this great scientist showed in creating the modern science of mechanics.

Incidentally you should note that the fact the ball rolls rather than falls does not invalidate either the distance-time relation or the 'circle-chord theorem'; this is because the rate of rotation increases in the same way as the rate of fall down the plane. The error only comes in when the extrapolation is made from the case of rolling down a vertical plane to that of falling freely.

If you already have some background knowledge of physics, particularly the idea of moments of inertia, you might like to consult Appendix 2 (Black) which goes into these points more fully.

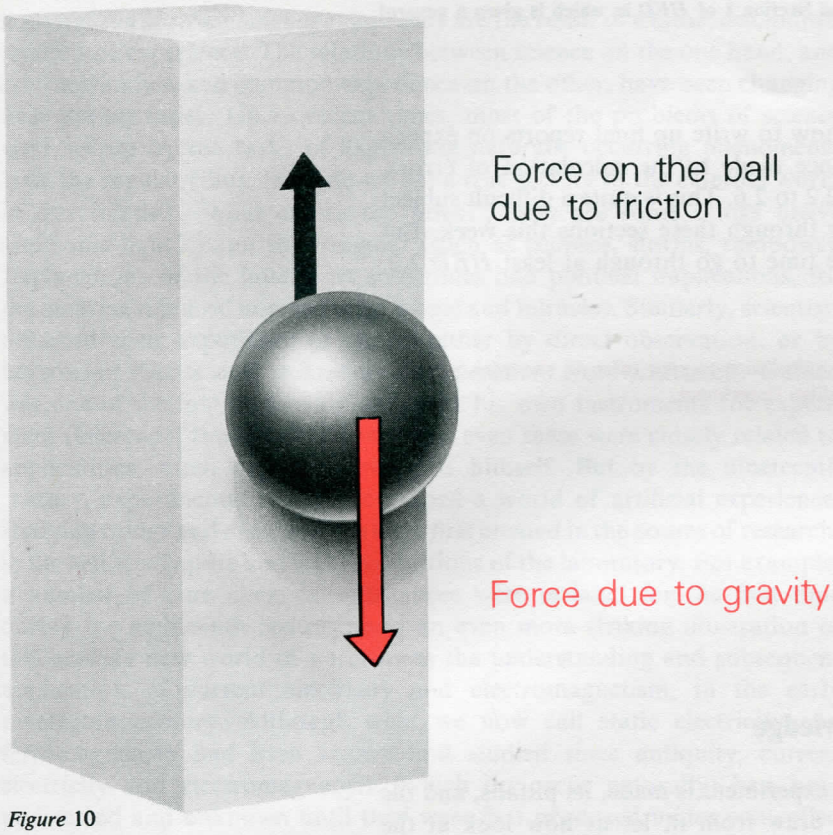


Figure 10

1.4.7 Some remarks about errors

Now that you have done an experiment, we want to draw your attention to the question of errors and limitations on the accuracy of experimental data.

You will have noticed that in repeatedly timing the ball rolling down a given distance at a given inclination, you did not always get the same value of the time.

What reasons can you give for this?

There are several possible reasons. No matter how smooth the surfaces of the ball and the inclined plane, there are always slight imperfections, specks of dust, etc. These will affect successive runs in slightly different ways. Though you may have thought of a very clever way of releasing the ball, it will not always have been released in *exactly* the same way. You will start and stop your watch in slightly different ways each time. These effects all give rise to random errors—sometimes they lead to an overestimate and sometimes to an underestimate of the time interval.

In addition there can be instrumental errors—your table top may not be perfectly flat, the watch may not respond immediately you depress the knob. There can also be 'personal errors'—you may have a systematic tendency to start the watch too early or too late compared with someone else doing this experiment (this is over and above the random fluctuations in timing). There are errors of bad technique—you probably notice that the reading on the watch differed depending on the angle at which you held the instrument—did you always hold the dial at the correct angle to take a reading? And how about extraneous influences like air resistance and draughts—in really accurate work these would have to be excluded.

These are some of the ways in which the data of this particular experiment could have been faulty.

We should now like you to read Section 1 of *HED* in which is given a general discussion of errors.

Read *HED* 2.14 which describes how to write up final reports on experiments. (You will find here reference made to the calculation of errors. This subject is dealt with in *HED* 2.2 to 2.6. This is quite a difficult subject and you need not work your way through these sections this week. But some time soon you must find the time to go through at least *HED* 2.2, 2.3, 2.4 and 2.6.)

Now write up in final form for submission to your tutor all the work you have done in connection with the Galileo experiment.

1.5 The nature of scientific knowledge

Having seen something of how an experiment is made, its pitfalls, and the sort of conclusions that you can draw from it, let us now look at the *nature* of scientific conclusions. What do we mean by a 'scientific' statement, compared with one in everyday language? Look first at these two statements concerning the phenomenon of falling bodies you have just been experimenting upon:

(A) Judging from their impact on landing, objects falling from a greater height fall for a longer time and arrive with a greater speed.

(B) For a body in unresisted fall, under constant acceleration g, $v = gt$, where v is the instantaneous velocity at time t after the instant of release from rest (do not worry about the definitions of these terms yet—they are defined in Unit 3).

What are the relations between these two sorts of statement?

Statement A derives from common experience and assumptions; it can be explained to someone without special training and can be shown by a simple experiment. Statement B contains several technical and mathematical terms whose meanings are not immediately obvious. It is precise, and its confirmation by experiment requires specialized equipment.

(a) List the terms in statement B not used in everyday language.

(b) In what ways does B tell a scientist more than A does? In what ways less?

(c) If you are trying to explain A to someone with no formal training in science, will it help to invoke B?

(d) How likely is it that B was derived purely by careful observation of the phenomenon described in A?

(a) Unfamiliar terms: "unresisted", "acceleration", "instantaneous", "velocity".

This example indicates that while much scientific knowledge has 'grown out of' everyday experience, the relations between the two are not simple and direct. Scientific statements have *precision* and *objectivity*; they use

a *specialized abstract language*; and they are the result of a close, disciplined analysis of experience. The relations between science on the one hand, and craft techniques and common experience on the other, have been changing over the centuries. Up to recent times, most of the problems of science were set up by the tasks of explaining naturally-occurring phenomena, both the regular (thus, 'how do we get a true picture of the external world in our minds?', 'what causes the tides?', 'why are some things heavy and some light?') and the irregular (such as eclipses, storms, rainbows). Explanations of the latter sort sometimes had political implications, for the answers replaced and excluded magic and miracles. Similarly, scientists obtained their experience of nature either by direct observation, or by borrowing results and techniques of experiment from craftsmen. Galileo was one of the first to develop or invent his own instruments for experiment (telescope, timing pendulum), and even these were closely related to applications, often ones he developed himself. But by the nineteenth century, experimental science developed a world of artificial experience, studying things and events which were first created in the course of research, in the artificially pure and stable conditions of the laboratory. For example, a number of pure chemical substances were isolated for the first time during the eighteenth century; and an even more striking illustration of this entirely new world of nature was the understanding and subsequent application of current electricity and electromagnetism, in the early nineteenth century. Although what we now call static electricity and ferromagnetism had been known and studied since antiquity, current electricity and electromagnetism (which do occur naturally) had been undetected and unknown until they were first produced under controlled conditions in the laboratory.

On the basis of its command of this new world of artificial nature, science could begin to penetrate industry on a large scale. The first application of current electricity and electromagnetism was in the telegraph, which suddenly enormously multiplied the speed of communication. Later in the century arose a totally new industry, providing a new source of usable energy—electricity. In the present century, all new industries depend on the application of scientific results; and even older craft industries (e.g. agriculture, textiles) are increasingly penetrated by scientific knowledge and a 'scientific' approach. Indeed, much of our present world of ordinary experience is now largely a product of science-based technology.

(b) B gives us a precise relationship rather than a general statement. It also makes the statement apply to any time, not just the moment of impact. But it mentions only time of fall, and not distance. Also it is limited to the case of "unresisted" fall.

(c) B would not be much use as an "explanation" of A since an understanding of B requires special knowledge of its technical terms.

(d) B could just possibly have been derived by observation; but to measure "instantaneous velocity" requires special equipment, and so B would require an *experiment* rather than merely looking at a falling body and commenting on what you observed.

1.6 The creative work of science

Most of the scientific facts used by scientists today, including those you will be learning in this course, have no obvious connection with craft practice or with ordinary experience. Indeed, as they are presented by teachers, it is often very hard to see where they came from, and by what processes. We will explain how the development of scientific knowledge separates it from its origins in creative enquiry; but first we consider the act of scientific creation itself.

The beginning of a scientific enquiry is a guess, a hunch, an idea of the way things might be. Unless the scientist is very unlucky, it will have some kernel of truth in it; but unless he is superhuman, it will also include errors. For example, an important error made by Galileo was in refining the common experience of falling things (statement A, p. 36) into the *law*† that freely falling bodies increase their speed in proportion to the *distance* fallen. This was an apparently simple scientific law, and only after years of work did he discover its falsity and obtain his equivalent of statement B, p. 36.

To describe the process by which the initial insight of a scientist is refined and corrected into a genuine result, we will use the idea of a 'problem' and to make this more precise, we will consider a set of examples of 'problems' (you are not expected to solve these here and now!):

(1) Given that $v=gt$ (as defined in statement B), calculate v if $g=9.8$ m s^{-2} and $t = 10$ s (m s^{-2} means metres per second per second.)

(2) Measure *instantaneous velocity* using a camera to take pictures every hundredth of a second to confirm by experiment that $v=gt$. *Velocity* is defined as the distance travelled in a given time in a given direction. In the present case, however, where the velocity of the body is continually changing, the distance travelled in a given time will only be a measure of the *average* velocity during that time interval. *Instantaneous velocity* is the velocity at a given instant and strictly speaking could be measured only if the time interval were vanishingly small.

(3) Assuming that air resistance is proportional to v, obtain a relation between v and t, for a body under constant acceleration (i.e. removing the requirement that the motion be unresisted).

(3a) Same as 3, but assuming air resistance to be proportional to the square of the instantaneous velocity, v^2.

(4) Decide by experiment whether the air resistance to a moving body is proportional to v, to v^2, or to some other power of v.

(5) Decide whether the air resistance to moving bodies is fundamentally different at high speeds and at low speeds?

What is involved in solving each of these problems? Problems 1 and 2 are straightforward exercises, one of calculation, the other of experiment. Problem 3 is a purely mathematical exercise. It could be put in the form of a 'differential equation' and then solved. Problem 3a looks similar but is not; for in this case the equations cannot be solved analytically, although approximate solutions are possible. Problem 4 is even more open-ended; the design of an experiment which would yield the appropriate data is a sophisticated exercise, and then one would need to 'fit a curve' to the data to estimate the correct 'index', if any one could be found. Problem 5 poses a very general question and requires a background in theory† and experiment (for example of the theory of 'turbulence' and 'turbulent flow') before it can be investigated in any detail.

Problems 1, 2 and 3 can be described as 'exercises'; by known techniques, one can achieve the known answer. When such problems are given to a student, their function is not to enable him to 'rediscover' known results, but to train him in the elementary skills of scientific work or to test his comprehension of the concepts involved. The other problems in the list are genuine scientific problems; there is no straightforward method using standard techniques whereby a reliable answer can be obtained. To the extent that any problem presented to the scientist is not a routine exercise, 'framing the problem', that is, asking the question appropriate to the problem and to the tools and apparatus that can be brought into use in answering it, is creative work. The process necessarily starts with a guess, a hunch, an insight or a speculation, sufficiently precise to indicate the path of development, but sufficiently vague to allow the idea of the nature of the problem to grow during the work.

1.6.1 Limitations on research

What determines the sort of hunch a scientist has, the sort of questions he asks of his material? Four factors, at least, are involved. One is his own capacity to design a crucial experiment, to think through the nature of the problem; his 'genius', if you like. A second is the equipment and apparatus he has at his disposal; their limitations determine the shape of the questions he can ask and hope to get answers. A good example of this is provided by the Galileo experiment you have just done. Other examples of these limitations in modern science will show up in later units, for instance, the limit of magnification of the optical and electron microscopes discussed in Units 2 and 14. A third is the prevailing spirit of the ideas of his fellow scientists. It is hard to think 'out of the rut' of the conventional wisdom of accepted ideas. Scientists who do so may even be unpopular with their colleagues, or their work not accepted for many years. But perhaps the most important overall factor is the dominant shape of the thought of the society in which the work is done, which is itself determined by the structure of that society. Thus in England of the nineteenth-century Industrial Revolution and of unrestrained industrial development, there was a tendency amongst the scientists to a very mechanistic view of the world. Mechanical models for force, heat and energy prevailed amongst the physicists, whilst terms such as the 'struggle for existence' and 'survival of the fittest' were in use amongst the biologists. By contrast, for example, the biological models of the late twentieth century make much more frequent use of terms such as 'community', 'interaction', 'feedback' and 'regulation', concepts more akin to today's prevailing ideology within Western society.

How a scientist gets the idea which produces the problem is a very complex question to which no very precise answers can, in fact, be given; no 'scientific method' will generate original problems. It also requires judgement to know whether a problem can be successfully investigated at a given point in the development of a science. Galileo himself was aware of problem 5 (p. 38), but he sensibly chose to ignore it in favour of problems that he could solve—those neglecting the effects of air resistance. Investigating significant new problems is hazardous work, for the chances of failure are high; the less that the paths to a solution are charted, the greater the pitfalls that lie in the way. As you study science, you will learn the names of great scientists who are credited with the discovery of basic facts and laws. These facts frequently seem so simple and obvious that it is hard to imagine why it took a great man to discover them. But this appearance may be deceptive; often the more simple and elementary the fact now appears, the more deep and difficult the work was which led to its first discovery.

1.7 Establishing scientific results

Conceiving an original scientific problem is an act of creation requiring the free play of the imagination; but once the scientist is involved in working on that problem, his inspiration must be supplemented by perspiration.

Without careful, systematic, self-critical work, his initial insight will remain half-baked. Thus, had Galileo contented himself with the plausible law that speed increases in proportion to distance, for which he could even provide a (fallacious) derivation from the law that distance is proportional to time-squared, he would never have achieved his successes in mechanics. This hard work is necessary because there is no standard method whereby any scientific problem can be solved, as on a computer. This is because no new scientific result can be 'proved' in the way that a student's routine exercise can be definitely solved, with one right answer out of the many wrong ones. A scientific result is achieved by the combination of reason and experience, and their union can never be simple or perfect. Scientific statements deal with abstract concepts. They are not expressed in the ordinary language of everyday use.

Do you recall an example of an abstract concept from within this Unit?

Yes, the concept of 'instantaneous velocity', statement B, p. 36.

Of course, the concepts used in science must in the last resort be related back to ordinary experience, so that scientists, who are only human, can reason with them (just as it has been observed that the strangest science-fiction monsters are composed of very ordinary parts). But a close relation to ordinary experience may actually be misleading. We speak freely of 'hot and cold', 'fast and slow', 'heavy and light', as opposed qualities, but if, for example, in describing the motion of bodies, we were to be restricted to words like 'fast and slow' we would find it impossible to master the science of mechanics. Indeed, this idea of 'opposing qualities' persisted until the time of Galileo, and caused trouble to him as well as to his contemporaries.

Quite soon in this course, you will encounter some concepts whose relation to ordinary experience is very indirect, and which can be studied only in the artificial conditions of the laboratory. (Recall the remarks on p. 37 about the development of a world of artificial experience by modern science). For example, the study in Unit 2 of the differing scales of length involves using 'electromagnetic radiation' of various wavelengths: radio waves, light, ultraviolet and X-rays. We can unify and explain these phenomena in terms of a theory of *electromagnetic radiation*, itself an effect of electric charge in accelerated motion. But what is 'electric charge'? We can *show* you some of the simpler effects of electric charge by asking you to do the experiment of rubbing a comb and seeing that it attracts bits of paper. But you still can't see or feel the charge itself, although you can see a spark or feel a shock. Thus we are really no further forward in answering the question: what it *is*? Scientific research does not attempt to answer such a question in a simple and direct way; rather, by building up knowledge of its properties, it approaches (but never reaches!) a 'true' understanding of its nature.

You will also see that scientific knowledge about any one concept is composed of statements about it in connection with *other* concepts. It might

seem that the whole process of definition and explanation goes around in a circle! In most practical cases, it does not. At any given point in the development of science, some concepts are better understood and more firmly related to experience; and these will be used as the foundation for those newly invented. However, in teaching the foundations of a subject, we find ourselves with an embarrassing freedom of choice; and we will not try to conceal this from you. For instance, as you will see in Unit 4, classical dynamics depends on 'mass' and 'force'. The definition of these has given trouble ever since Newton's first attempt in his classic book *Principia* (1687). One can only measure mass through some effect *involving* forces such as collision, or gravitational attraction. But then if one considers force as a cause of motion, one can measure it only in terms of the velocity it imparts to a given mass! At this point, philosophers of science can assist the advancement of scientific knowledge; for they analyse the concepts that scientists develop and use in a practical way, and examine how they might be clearly defined and consistently related among themselves.

1.7.1 The design of experiments

It would be comforting to think that all these difficulties with concepts are relatively unimportant, since science depends basically on 'facts' which are results derived from experiment or observation of the natural world. But here too the work of science requires skill and judgment, and is open to error. For whatever the source of the data the scientist uses, it is inevitably imprecise (see *HED*, section 1) and can only approximate to answering the questions which the scientist is really asking. Also the experiment itself must be designed for the testing of those effects that can be studied with the available tools. These will usually be very special ones (as the 'circle-chord theorem' in the Galileo experiment) which are not found in direct observation of nature. To the extent that the conditions of the experiment are different from those 'outside', there will be difficulties and pitfalls in applying the conclusions of the experimental research to natural phenomena themselves. You have already seen one such classic pitfall, in the very simple case of the rolling of bodies down an inclined plane; imagine how much more difficult is the job of extrapolating from an experimental situation in a complex physical science, to say nothing of the biological, or earth sciences. We can illustrate the inherent difficulties and limitations of experiment, by considering problem 2, p. 38: measuring instantaneous velocity by a stroboscopic camera (one that takes pictures by light-flash every given fraction of a second).

> **Why have pictures every hundredth of a second? Why not have a simpler apparatus, and take them every tenth of a second? Or should we rather do the job properly, and have pictures taken every nanosecond (10^{-9}s)? (N.B. You will learn how such short time intervals can be measured in Unit 2.)**

The answer is that for some purposes, a tenth of a second is not quite good enough, a hundredth of a second is adequate, whereas a nanosecond gives too much detail. The time-intervals must be small enough so that the *average velocity* between them is a good approximation to measuring the *instantaneous velocity* at any point of time in the interval. Also, there must be enough estimates of this average velocity for a test of any hypothesis about the behaviour of the falling body. If the body falls only 1 metre, the total time of fall is less than half a second, so tenth-second photographs would give only three average velocities, but a fall of 5 metres would give about a second for the fall, and enable ten estimates of average velocities. On the other hand, if one is concerned to see the effects of air resistance, turbulence, etc. at very low velocities, then one might want a very closely

spaced sequence of photographs, for which even hundredth-second intervals might be insufficient.

Does this experiment measure 'instantaneous velocity'?

Of course not! Even a nanosecond interval would not do that. But if the thing can't be measured then what is it?

Even when we take a series of photographs as described above, they will be slightly blurred; there will be a distortion near the edge of the frame, and there will be systematic and random errors in the timing of the photographs. For the purposes of the sort of measurements that we are trying to make here these effects don't matter; but this is because the progress of science, made possible in this case by Galileo's work, results in the technological development of tools such as clocks and cameras, which render a repetition today of his original experiment apparently easy and trivial. However, the powerful and complex tools used in contemporary science, such as the high-energy accelerators of elementary particle physics, are designed for experiments done at the very limit of their own capabilities. There we find the same difficulties in the production and interpretation of experimental data that Galileo had with his primitive equipment.

Highly accurate experiments with bodies falling through the air at different speeds, would not 'confirm' the relation $v = gt$; they would show that it is false. The experimental data will fit the predicted curve, only when speeds are not too great and measurements not too precise. This deviation of experiment from theory is due to air resistance (and also partly to the inevitable experimental error that remains even in 'highly accurate' experiments). But study of the effects of air resistance presents problems of very great difficulty. Most of the mathematical equations derived from the physical data cannot be solved in any neat way; the measurement of air resistance is itself a difficult procedure; and the interpretation of experimental data is complex.

1.7.2 Models

A student doing a 'set' experimental exercise can obtain his data and derive a result that he knows agrees with that quoted in a text book. But we have seen that for the scientist who is handling imprecise data in the hope of getting evidence for a conclusion about artificial concepts, life cannot be so simple. For the scientist cannot 'prove' a new result in his research; at most he can establish it by an argument that is 'probable' and 'reliable'. But the patterns of argument in science are not those of ordinary persuasive or plausible reasoning, like those of the politician or advertiser, for instance. They depend not only on the marshalling of experimental observations, but on the analysis of those observations according to a set of logical procedures that are universally acceptable to scientists working in that particular field. The type of procedure used differs somewhat from one scientific field to another, depending, for instance, on whether it is predominantly an experimental field (like today's biochemistry, for example) or a theoretical one, where experimental data is scarce but theoretical ideas are plentiful, like cosmology, at least until recent years. In the former case the argument might proceed by the presentation of a new experimental result which fits, or fails to fit, a formerly accepted theory. In the latter, much of the scientist's work may go into the making of a *model*†: a conceptual or mathematical structure whose properties can be systematically explored, in the hope that some of these will agree with experiment, and so provide clues to the causes of

the phenomena. The relation between the model and the real things and processes is one of *analogy*: a similarity between important features on both sides.

Can you think of some examples of such models or analogies?

Some you will meet later in the course include billiard balls for models of molecules of a gas (Unit 5); waves in water as a model for sound or light waves (Unit 2); a sheet of rubber deformed by a stick as a model of a field of force (Unit 2); a pump as a model for the heart (Unit 18). You will find more examples in succeeding course units.

In a sense all scientific theories are models, based on analogy; the mathematical point whose motion is described by the relation $v = gt$ is a *model* for the *real* falling body. But in cases such as this, the similarities are so strong that for many purposes we can consider the equation as *really* describing the behaviour of the falling body itself. When the scientist operates in this way he hopes to have something stronger than a model. For, if his conclusions *explain* the phenomena in some sense, he can consider himself as having achieved a *theory*. You will soon see an example of such a 'theory', in the explanation of light as consisting of electromagnetic radiation (Unit 2); another is Newton's theory of universal gravitation (Unit 4).

1.7.3 Hypotheses

On the other hand, when the work on a problem involves some interaction with the real world, through experiment or observation, the role of *hypotheses*† becomes crucial. An hypothesis is a statement about the things being studied, put forward not as a conclusion to be accepted and developed, but only as a tentative conjecture, serving as a guide to investigation, rather like the 'hypotheses' that we suggested were made by rats as to which doors had food behind them. Observation and experiment must be done in a systematic and disciplined way; good scientists do not simply 'go out and collect' data. Although chance discoveries may start off a problem, the data which is eventually used must be produced methodically, so that it is as free as possible from human error in its compiling, and also protected against errors of interpretation. A common way of tackling a problem is to design the experiment so that it *tests an hypothesis*. In doing this, the scientist must state exactly what the experiment is about; he must organize the production and recording of data into a particular routine; he can be clear about the sorts of data that confirm or refute the hypothesis (or leave it undecided); and he can more easily locate errors in the interpretation of the data.

The way in which an experiment is designed to test an hypothesis must be carefully planned if correct conclusions are to be drawn. This is especially true when the data cannot be interpreted at a glance, either as they stand, or when plotted on a graph. For example, when the scientist wants to know whether two effects are independent of each other, or are related so that when one occurs the other also tends to, then it is necessary to analyse the results mathematically and statistically. These statistical techniques are very powerful, so much so that quite unrelated sets of data can, with desire and ingenuity, be shown to be 'correlated'. Examples of such false or misleading correlations come most frequently from the medical sciences, although they are certainly not unknown elsewhere. One of the most striking in recent years was produced in a study of the rapid increase in the occurrence of deaths due to the disease of coronary thrombosis, when it was shown that the only factor examined whose increase in the community paralleled that of the disease was the issue of television licences. It could be argued to 'follow' from this that ownership of a television set—or a licence for one at any rate—was a factor in causing the deaths by coronary thrombosis. However, no one would regard such a correlation as being of much use in helping find an explanation or a cure for coronary thrombosis!

To guard against such pitfalls, the founders of modern statistical methods devised principles and rules for such work; foremost among these is that every experiment be designed to test a particular hypothesis. From the results of the experiments a statement of statistical probability can be made, the probability that the hypothesis in question is true. The scientist must then use his judgement, aided by the experience of his field, to decide whether a particular level of probability or truth (or falsity) is good enough. For some purposes, 95 per cent (odds of 20 to 1) is good enough; but for others he cannot be satisfied with less than 99 per cent (100 to 1). If you are not clear about such aspects of probability, look at the examples given in Appendix 1 (Red) (p. 49).

1.7.4 Deductive and inductive inference

Scientific problems vary enormously in their degree of theoretical complexity and in their dependence on data derived from experience; and so the patterns of argument used in them, and the methods of work, will be correspondingly various. Depending on his particular experience of research, one scientist may think of 'scientific method' as the construction of models and theories, and will be concerned with the relation of theoretical entities to the real world; while another will see 'scientific method' as a series of tests of simple hypotheses converging towards a true answer to the problem. But in any scientific problem, the argument from which the final conclusions are drawn will include several sorts of inferences. One sort is *deductive*, proceeding by logical steps from assumptions to *conclusions*, as in a mathematical proof, e.g. that in a right-angled triangle, the square on the hypotenuse is equal to the sum of the squares on the two other sides. An example of this is provided by the use of geometry and trigonometry in the circle-chord theorem of the Galileo experiment. A second is *inductive*, generalizing from the properties of a sample to those of the large group from which it comes. (Such a large group is called a *class*.) For instance, it is an inductive inference when we say that *all* one-celled organisms avoid sulphuric acid, based on a study of only a few, or, that *all* rats can make hypotheses or *all* birds will shun black and yellow striped insects. You used inductive inference in generalizing the results of your experiments on rolling a ball down a table to apply to the use of *any* sphere rolling down *any* inclined plane. Some inferences are *probabilistic*, giving an estimate of the confidence that can be attached to a particular statement, as in the examples above, or in averaging from the spread of readings you obtained in the Galileo experiment. Finally, *analogical* inference, arguing by extension of accepted similarities, is used in connection with models, as when we say that *because* the molecules behave like billiard balls in such-and-such a situation therefore they must have such-and-such other properties. You used this type of inference when you assumed in the circle-chord theorem, that the rolling ball could be represented by a mathematical point, that the behaviour of the rolling ball could be extended to that of a falling body, and finally, that a large, irregular body could be treated as if it were a small sphere.

During the course of the year's work you will find many different examples of the use of these types of inference. Note them, for at the end of the year you should be able to quote further examples derived from subsequent course units.

What emerges from a scientist's work on a problem is not yet a 'fact'†. Although the scientist can present his conclusions *as if* it follows necessarily from hard experimental data, he and his colleagues know that it is liable to error at a multitude of points: at the production of this data, when errors in making observations can occur, in relating the data to the concepts being studied, and in the structure of argument itself. If he has worked properly, the scientist will have guarded against the inevitable tendencies to self-deception; he will have made his data as precise and

objective, and his argument as rigorous as possible. But perfection is impossible, and the result must be further tested before it can be considered as a 'fact'.

It is just because the results of scientific research cannot be perfect that the work itself is very demanding. Conceiving a genuinely new scientific problem requires an active curiosity; and following it through requires courage (for it may be an unpopular or risky line, or even simply wrong), as well as stamina. Since results can never be perfect, the scientist must constantly be exercising his judgement: is this part of the work *good enough*? There is no simple testing gauge for scientific results; and the people who will test a scientist's work are scientists like himself. If all the scientists in a field are concerned to get by with only a minimum of effort, really first-class work will become increasingly rare, and the field will become mediocre, boring and stagnant. This is why a particular sort of ethical commitment is necessary for worthwhile scientific work; and it requires leadership and morale for its maintenance.

Section 8

1.8 The achievement of scientific knowledge

Once a scientific result has been achieved it enters on a second stage of development, where the processes are *social* rather than *individual*. These may be described as *testing*, *selection* and *transformation*. Only a very small proportion of scientific results survive these processes; and it is these that make up the body of permanent results, and genuine scientific knowledge. The testing of a new result is done in two stages: first, it is scrutinized by other independent scientists before it is published in a recognized scientific journal; and it may be rejected or returned for modification. Publication as a scientific paper is the recognized form of announcing scientific results. This form of announcing results to the world emerged in the seventeenth century (before then scientists had put their results into letters to other scientists or had published them as books). Today, there are upwards of some 33 000 scientific journals in the world, often published weekly, and even the *abstracts* of articles in one field, such as chemistry, can occupy 15 000 pages a year.

If the result is published, it is subjected to selection: does anyone take notice of it? If not, it dies there and then; and this is the fate of a large proportion of the published papers in any field. The reasons for this are numerous. They include the following: first, although the work is competent, the result leads nowhere and neither suggests worthwhile new problems nor contributes useful information or techniques. An example would be a study of the number of times the letter 's' appears in the play *Hamlet* compared with *Macbeth*. Secondly, the paper uses methods that are so new and unfamiliar that no one understands it; or the results seem so strange that the work is dismissed as that of a crank. The fate of such unorthodox papers depends very much on the practices of the field to which it is offered; if it is dominated by strong men with strong ideas of what is correct and worthwhile, there is a strong risk of unconventional work being ignored. This seems to have been more prevalent in the nineteenth century than it is today. Thirdly, shortly after the appearance of the paper, another one, doing similar work more efficiently, appears. An example would be a paper describing a method for the purification of a chemical which takes 24 hours and has a 10 per cent yield of 90 per cent purity, which is followed by a paper giving a different method which takes only 10 hours, has a 70 per cent yield and a 99 per cent purity. Fourthly, when other scientists try the methods, they do not work well enough to be

worth applying to other problems, and lastly, the result is subsequently found to be wrong! In this case it might be used and cited briefly but will then quickly be forgotten.

1.8.1 Facts and laws

We notice that the various reasons for the non-selection of a result are concerned only with its quality; the personality, race or politics of the author is irrelevant to the assessment. This ideal is not always easy to adhere to in practice, especially when scientists are subjected to outside pressures from politicians, or when a field is split into antagonistic rival schools. But this strict ethic of scientists, a special sort of 'fair play', is just as essential to the health of science as is the individual's commitment to doing good work. If against all these obstacles a result is selected, it is then tested again through use, or the result itself may suggest further problems. If it turns out that it did not live up to its promise of usefulness, then it is quietly dropped; and it is then forgotten and dead. But if it stays alive, then quite soon it is subjected to a process of *transformation*. For as science advances, earlier results seem crude and obsolete; they must be translated into the new language. This can happen relatively rapidly, when there is a 'scientific revolution' replacing one set of concepts by another, but more commonly, it is a gradual, continuous process (the best-known example of such a major transition is, perhaps, the replacement of Newtonian gravitational theory by those of relativity by Poincaré and Einstein at the beginning of the twentieth century). If the original result cannot be translated into a useful new form, it dies at this point; but if it survives such a transformation, we can be confident that it does say something real about the world of nature; and then it is called a fact.*

A result which becomes an accepted fact need not be restricted to a statement of the properties of some particular thing (for example, the melting point of a metal). Some facts can be considered as *laws* which give a unified description of a whole range of phenomena. Some such laws can be demonstrated by simple experiments (as Galileo's law) while others have only a very indirect connection with experience and are accepted because of their power for explanation and prediction, and their coherence in unifying a mass of apparently unrelated facts. Examples of this sort are the First and Second Laws of Thermodynamics (which you will come t in Unit 5). Even such facts and laws are subject to the ravages of time. They may be revealed as crude and eventually uninteresting approximations, or the whole field of study in which they were developed may sink into oblivion.

Examples of such fields are difficult to give, because by definition they are almost forgotten (except by historians of science), but they include for instance the ramifications of the theory of the stars, Sun and Moon as moving in 'celestial spheres' around the Earth—what was called Ptolemaic astronomy. This was destroyed by the observations and theories of Copernicus, Galileo, Kepler and finally of Newton in the sixteenth and seventeenth centuries. A similar episode occurred in the nineteenth century when, because light was believed to move in waves, and waves needed to be 'in' some medium, a fluid called 'aether' which pervaded space, was postulated. You will learn more about wave motion in Units 2, 22, 28 and 29 and you will see how a totally new explanation, Einstein's theory of special relativity, made all the earlier theories about the medium of

** We should perhaps point out that some philosophers would dispute whether anything can ever be claimed to be a 'fact'. But certainly scientists at least are prepared to accept various results as facts.*

wave propagation irrelevant. A final example is provided by the attempts made by seventeenth-, eighteenth- and nineteenth-century researchers to transfuse blood between animals and man. The attempt failed, because not enough was known about the problems of immune reaction, 'rejection', and other similar bodily defence mechanisms. These could only be studied meaningfully, and the problems of transfusion overcome, in the twentieth century, as a result of advances in biology.

Facts remain in the forefront of research only for a limited time and then pass into the body of standard information used in the basic teaching of science. An example of such transformation is provided by Galileo's law. He first stated it as a property of 'naturally accelerated motion'. When Isaac Newton introduced the idea of 'force' into mechanics some fifty years after Galileo's death, the law became a 'theorem' about motions produced by certain types of force. Later, in the nineteenth century, when the law of the 'conservation of energy' (Unit 4) was applied to mechanics, the theorem was interpreted in terms of a *relation* between kinetic and potential energy in a particular system. Finally, it became a model 'experiment' used in teaching.

In this last use, such facts are shaped by the needs of teaching situations, and will retain only the faintest resemblance to their originals. What was originally a deep, difficult and perhaps obscure result, achieved in the context of advanced research, is transformed into something simple and clear, designed for easy comprehension and manipulation. Doing the experiment has shown you this in the case of Galileo's law, and you will see it again in much of the rest of this course. Thus the materials of any course in science, including this one, are the product of a lengthy process of evolution—and indeed are still evolving.

* * *

At this point you have reached the end of the main text material of Unit 1 in the Foundation Course in science.

Can you summarize here what you think are the main points and sequence of arguments in this Unit, in about half a page of writing? As a guide, look again at the list of objectives on page 8.

When you have done so—and not until—turn over the page and compare your summary with that we have made and with the conceptual diagram on pages 6 and 7.

1.9 Summary

This Unit, the first in the Foundation Course in science, starts by discussing the relationships between science, technology, and society. It shows how much of modern life is shaped by science and technology. This represents section 1 of the Unit. In trying to understand the nature of science, the question is asked: how did it arise, and is man's science unique? Section 2 tries to answer this by looking at the evolutionary pathway which leads to man; it shows that activities which are in some sense 'scientific' are apparent in the behaviour of many types of non-human animal. The evolutionary pathway which has led to man is one of increasing power of 'brain' and of capacity for explorative or creative behaviour. With man, these properties reach a high point which, even in primitive societies, can be seen to emerge from craft, religion and magic and extend towards what we now call science. The dawn of science is discussed in section 3. But science and scientific procedures are, as shown in sections 4 and 5, more than just craft knowledge, they are precise and objective and represent both knowledge *and* power. They require *creative work* involving the setting and solving of new problems. This is discussed, using as an example a seemingly simple experiment performed by the seventeenth-century Italian scientist, Galileo. Section 4 goes through this experiment in detail. Sections 6 and 7 ask the question 'how is a scientific result established?' and show how the procedures of making experiments and testing *hypotheses*, *laws* and *theories* have developed. Section 8 shows how results, when established, pass into the scientific literature and become 'transformed' by use into facts, knowledge—and teaching material.

Book List

Preparatory reading

J. D. Bernal, *Science in History*. Penguin, 1969, 4 volumes.

Set books

J. R. Ravetz, *The Roots of Present-day Science*. Open University, 1971.
H. Rose and S. Rose, *Science and Society*. Penguin (Chapters 1–6), 1970.

Background reading

T. S. Kuhn, *The Structure of Scientific Revolutions*. Chicago UP, 1962. (2nd edition, 1970.)
P. B. Medawar, *The Art of the Soluble*. Penguin, 1969.
J. M. Ziman, *Public Knowledge*. CUP, 1968.

Appendix 1 (Red)

Some examples of probability

Jones said to Smith, 'I am a very lucky man. My telephone number is 65–6237. My chances of getting this particular number were 1 in a million.'

What was the fallacy in Jones' reasoning?

For Jones to consider himself truly lucky, he would have needed *first* to make an hypothesis about the probability of obtaining a number in some class or classes (for instance, all digits the same, digits in sequence, a good poker hand etc.) and *then* see if such a number was assigned to him.

An American businessman got a phobia about air travel; he was afraid that he would be killed by a crash resulting from a bomb smuggled on board his plane. His firm learned from their insurance brokers that the odds against a plane carrying a smuggled bomb were 10 000 to 1. But the man was not satisfied. So he was advised to smuggle a bomb himself, for the odds against a plane carrying *two* bombs are 100 000 000 to 1.

This made him feel much more secure.

Was he deluding himself?

Yes, the probability quoted for two bombs only applies to the situation where there are two *independent* chance events. His own bomb was not a "chance event" in this sense. The problem could be put thus:
If a plane carries one smuggled bomb, the probability of its carrying a second one is the same as it was for the first: 1 in 10 000. So the man, carrying his bomb, was no safer than before.

A schoolboy did the experiment of 'testing the hypothesis' that $v = gt$, where $g = 9.80$ m s^{-2}. He reported his results in the following table:

v (m s^{-1})	9.80	19.60	39.20	78.40
t (s)	1	2	4	8

He was mortified when his lab report was returned with the note: 'Failed; *Do* the experiment next time'.

Why was the teacher dissatisfied with this perfect confirmation of the hypothesis?

It is highly improbable that in a real experiment the results obtained would fit the "hypothesis" so perfectly and without errors due to the sort of measurement difficulties we discussed in this unit. Therefore the teacher could be reasonably sure that the student had cooked the data or perhaps had never done the experiment.

(You have been warned!)

Theory of the Galileo experiment

(a) Proof of the 'circle-chord' theorem

Refer to Figure 11

It is easy to calculate the time the ball takes to roll from C to B from the principle of conservation of energy. Suppose the ball has mass m. When it is at C it will have potential energy which is greater than it would have at B by the amount mgh, i.e. the work done against the gravitational force mg, in lifting the ball through a height h, where g is the acceleration due to gravity. If it starts from rest at C and by the time it has rolled down to B, has a velocity v, its kinetic energy (neglecting for the moment the energy that goes into making the ball rotate) will be $\frac{1}{2}mv^2$.

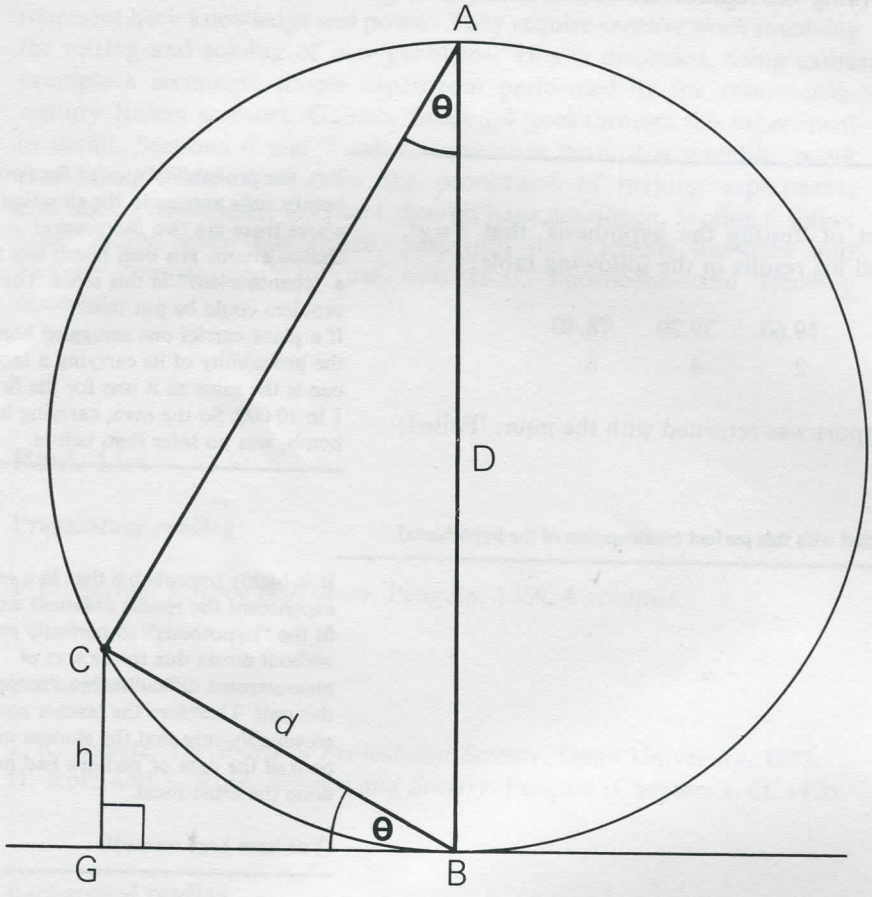


Figure 11

Then, by the principle of conservation of energy, the gain in kinetic energy is equal to the loss in potential energy.

That is $\frac{1}{2}mv^2 = mgh$

or $v^2 = 2gh$

$v = \sqrt{2gh}$

Now, since the ball is acted on by a constant force, it accelerates uniformly and so the *average* velocity, $\bar{v}$, is just half the final velocity v.

Thus, $\bar{v} = \frac{1}{2}v = \frac{1}{2}\sqrt{2gh} = \sqrt{gh/2}$

The time to roll the distance d is then

$$t = d/\bar{v} = d/\sqrt{gh/2} = d\sqrt{2/gh}$$

But $h = d \sin \theta$ (see Fig. 11)

So $t = d\sqrt{2/gd \sin \theta} = \sqrt{2d/g \sin \theta}$ ..... (1)

and $t^2 = 2d/g \sin \theta$ ..... (2)

Now look at Figure 11 again. ACB is a right angle, as is ABG. It follows that the angle CAB is also θ and that

$$\sin \theta = \text{CB/AB} = d/D.$$

Substituting for $\sin \theta$ in equation 2

$$t^2 = 2D/g \quad \text{..... (3)}$$

Thus t has the same value for any chord of the circle of diameter D

This is Galileo's 'circle-chord theorem'.

Note that if $t = 1$, then, from equation 3,

$$D = g/2$$

So, *in the absence of rotation*, the diameter of the 'one-second circle' should be 4.9 m, since g is about 9.8 m s^{-2}.

(b) The effect of rotation

We now take into account the energy that goes into making the ball rotate.

The *rotational* kinetic energy of a body rotating with angular velocity ω is $\frac{1}{2}I\omega^2$, where I is its moment of inertia about the axis of rotation.

For a solid sphere, the moment of inertia about a diameter is $\frac{2}{5}mr^2$, where r is the radius.

So the rotational kinetic energy of a rolling spherical ball is

$$\frac{1}{2}I\omega^2 = \frac{1}{2} \cdot \frac{2}{5}mr^2 \cdot \omega^2$$

$$= \frac{1}{5}mv^2,$$

where v is the linear velocity of the ball down the inclined plane. (Note that $r\omega$ is the linear velocity of the point of contact between the rotating ball and the inclined plane, relative to the centre of the ball. As this point of contact is at rest relative to the inclined plane, and the centre of the ball is moving with a velocity v relative to the plane, it follows that $v = r\omega$.)

The *total* kinetic energy of the rolling ball is thus

$$\frac{1}{2}mv^2 + \frac{1}{5}mv^2$$

and this must be equated with the potential energy mgh.

So $(\frac{1}{2}+\frac{1}{5})mv^2 = mgh,$

or $\frac{7}{10}mv^2 = mgh$

What effect will this have on the time of descent?

If you follow through the calculation as before, you will find

$$t = \sqrt{2.8d/g \sin\theta} \quad \text{..... (4)}$$

This is $\sqrt{1.4}$ times longer than the time given by equation 1, in which we neglected rotation.

So, the diameter of the 'one-second circle' you should obtain from your experiment is

$$D = g/2.8 = 3.5\text{m}.$$

You may like to verify, as an exercise, that if the moment of inertia of any body about an axis of rotation is I, and that body is rolled down an inclined plane in such a way as to make it rotate about that axis, then the time of descent will be

$$t = \sqrt{2(1+\alpha)d/g \sin\theta} \quad \text{..... (5)}$$

where the symbols have the same meaning as before, and

$$\alpha = I/mr^2,$$

r being the radius of the rolling body.
As a check, note that for a solid sphere,

$$I = \tfrac{2}{5}mr^2,$$

so $$\alpha = \tfrac{2}{5} = 0.4,$$

and equation 5 becomes identical with equation 4.

You can use equation 5 to explain the results of the 'cylinder race' in the television programme of this Unit.

To do this, you will need to know that the moment of inertia of a hollow cylinder about its axis is

$$I = \tfrac{1}{2}m(R^2 + r^2)$$

where R is the outer radius and r the inner radius of the cylinder.

(See the post-broadcast section of the Broadcast Notes for this Unit's television programme for further information.)

Appendix 3

Glossary

ASDIC A device for detecting submarines developed in the Second World War. The word is made up of the initials of Allied Submarine Detection Investigation Committee.

FACT An assertion that is accepted as correct in the context of its use. The results of research reach the status of being facts only after the social processes of selection, testing and transformation in subsequent research and teaching.

FACTOR To increase the size of a thing by a *factor* of ten means to multiply the size by ten.

HYPOTHESIS An assertion put forward tentatively, as a guide to research rather than as a firm statement of fact.

LAW (scientific) An assertion that unifies and also serves for the explanation of a collection of facts; and is generally held to be correct within the limits of its statement.

MODEL A system, either of concepts or a physical device, whose structure and properties have important similarities to that of the state of affairs being studied. The relation between the model and its object is one of *analogy*: a similarity in some respects but not in others. A model may be advanced as an hypothesis, or used as part of a theory.

PANTOGRAPH An instrument for copying plans, diagrams, etc., on any scale.

THEORY An assertion which serves to *explain* a state of affairs; it is not claimed to be completely correct or adequate, but it is held more strongly than a mere hypothesis.

Section 1.1

Question 1 (*Objective* 2)

Tick to show by what factors each of the following human capabilities appear to have increased approximately over the last century?

	A 10^1	B 10^2	C 10^3	D 10^4	E 10^5	F 10^6	G 10^7	H 10^8	I 10^9	J 10^{10}
1 Communication	☐	☐	☐	☐	☐	☐	✓	☐	☐	☐
2 Speed of travel	☐	✓	☐	☐	☐	☐	☐	☐	☐	☐
3 Data handling	☐	☐	☐	☐	☐	✓	☐	☐	☐	☐
4 Tapping energy sources	☐	☐	✓	☐	☐	☐	☐	☐	☐	☐
5 Destruction by explosives	☐	☐	☐	☐	☐	✓	☐	☐	☐	☐

Question 2 (*Objective* 3)

Place a tick in the boxes opposite those statements below which you recognize as examples of modifications caused, or being caused, to modern society by science and technology.

1	The increasing incidence of cremation	☐
2	The payment of invoices through GIRO	✓
3	The decreasing economic viability of cinemas	○
4	The widespread wearing of mini skirts in 1970	☐
5	The increase in pollution of the environment	✓
6	The change in laws relating to male homosexuality	☐
7	The absence of fireplaces in bedrooms of new houses	○
8	The existence of the concept of the 'Third World'	☐
9	The layman's increase in awareness in outline, of world affairs	○
10	The existence of the Open University	○
11	The change in attitude to women seen smoking in public	☐
12	The increased tolerance of nudity in plays and films	☐
13	The maintenance of stockpiles of nuclear weapons	✓
14	The decrease in child mortality in Western Europe in the last 20 years	✓

Question 3 (*Objective* 4)

> So science and technology must be seen as interdependent activities, discovery precedes invention, and invention in turn makes possible discovery. . . (Unit 1, p. 11.)

Think about this in terms of:

(a) discoveries and inventions suggesting new avenues for exploration not thought of before, and
(b) experiments or explorations conceived of, but impossible in practice until some technological advance, perhaps many years later, and in some totally different context, makes them practicable.

Illustrate your thinking by producing three paragraphs, each of which shows an example of the reciprocal interaction between science and technology. Either make up your own material entirely, or use terms from List A and List B below.

List A

Results of 'pure' and 'basic' research: 'science'

concepts or discoveries

1 radioactivity
2 mascons
3 the abnormal number of genes in Mongol children
4 cathode rays
5 age of the earth
6 expanding universe
7 atoms and nuclei
8 radio waves
9 polymerization
10 hormones
11 theories of optics
12 cosmic rays
13 viruses
14 air as a substance
15 blood groups
16 recent ideas about Stonehenge
17 electromagnetism
18 terrestrial gravitation
19 waves travelling through the earth or water
20 superconductivity
21 X-rays
22 momentum and inertia
23 light sensitivity of some chemicals
24 antibodies and antigens

List B

Results of 'applied' and 'developmental' research: 'technology'

technical achievements

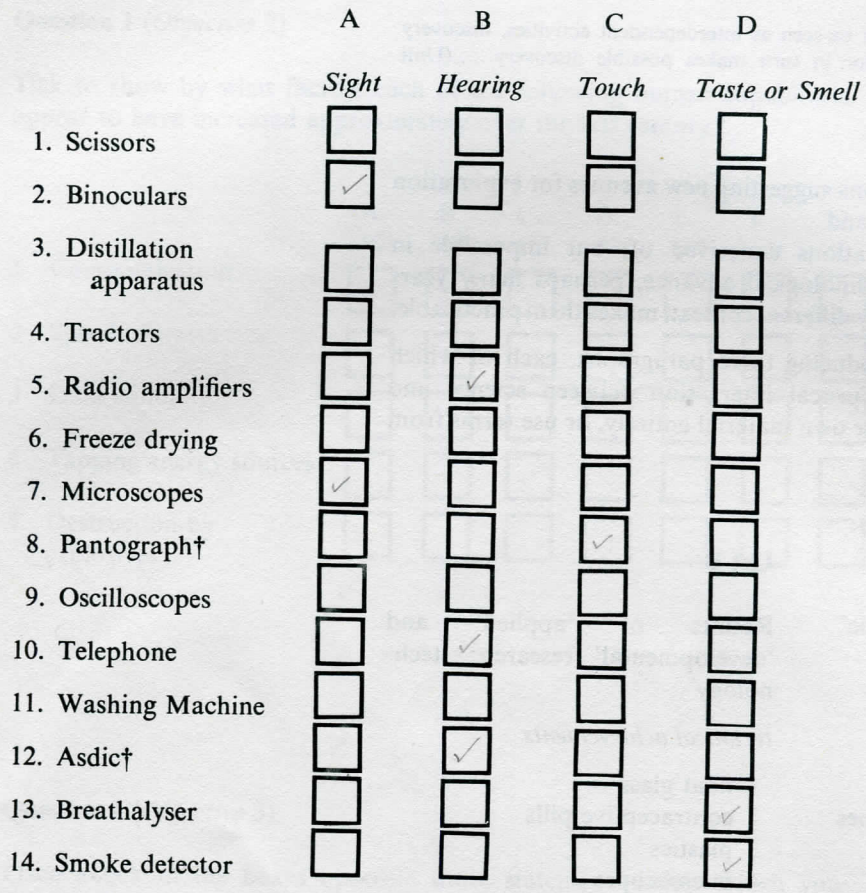
1 float glass
2 contraceptive pills
3 plastics
4 microscopes
5 high speed photography
6 radiotherapy
7 atom bombs
8 computers
9 spectroscopes
10 tissue transplants
11 seismographs
12 blood transfusion
13 television
14 telescopes
15 refrigeration
16 Geiger counters
17 frozen foodstuffs
18 vacuum pumps
19 rocket fuels
20 artificial satellites and space probes
21 ultracentrifugation
22 radar
23 treatment of diabetes
24 petrol engines
25 helium balloons

Note: You are *not* expected to be familiar with all the concepts, discoveries and technical achievements listed above!

Section 1.2.1

Question 4 (*Objective* 5)

Tick to show which of the five senses, if any, is extended by which of the 14 devices or techniques listed below.



	A *Sight*	B *Hearing*	C *Touch*	D *Taste or Smell*
1. Scissors	☐	☐	☐	☐
2. Binoculars	✓	☐	☐	☐
3. Distillation apparatus	☐	☐	☐	☐
4. Tractors	☐	☐	☐	☐
5. Radio amplifiers	☐	✓	☐	☐
6. Freeze drying	☐	☐	☐	☐
7. Microscopes	✓	☐	☐	☐
8. Pantograph†	☐	☐	✓	☐
9. Oscilloscopes	☐	☐	☐	☐
10. Telephone	☐	✓	☐	☐
11. Washing Machine	☐	☐	☐	☐
12. Asdic†	☐	✓	☐	☐
13. Breathalyser	☐	☐	☐	✓
14. Smoke detector	☐	☐	☐	✓

Sections 1.2.3, 1.2.4

Question 5 (*Objective* 6)

A. Place a tick in the box opposite any statement which suggests strongly an action that depends on brain memory.

1	An animal consistently observed to stalk its prey from down-wind direction	☐
2	A chick reared by hand in isolation will open its beak when shown a good model of its parent's beak	☐
3	An animal jumping on receipt of an electric shock	☐
4	The successful training and use of tracker dogs	✓
5	An animal born blind seeks the maternal teat without help from the mother	☐
6	A man exhibiting fright on being taken into a torture chamber	✓
7	A breech baby may extend its leg upon sharp stimulation of the sole, contrary to the normal reaction of non-breech babies	☐
8	Birds reared in isolation have been observed to attempt nest-building	☐

9 A person putting up an umbrella upon seeing a black sky and hearing thunder ☑

10 A person taking his finger away quickly after placing it on a hot object ☐

B. You should now be able to separate the statements you have ticked into two groups because they represent behaviour patterns suggesting two different origins. Write, in the boxes below, the two origins and the statement numbers relevant to each of them:

	First Group	Second Group
Origin		
Statement Number		

Section 1.2

Question 6 (*Objective* 9)

Two sets of statements are given below. Those in List A correspond to characteristic evolutionary features of man; those in List B to capabilities of man. Sort out three features from List A which distinguish man from all other organisms and place the appropriate numbers one in each of the three small boxes below. From List B select numbers which exemplify man exploiting the features selected from List A and place the appropriate List B numbers in the larger boxes, thus associating them with a specific feature (a number from List B may be used more than once).

List A

Characteristic evolutionary features of man

1 Growth of hair on his head
2 Communications by means of language
3 Brain/hand co-ordination
4 Possession of hand with freely moving thumb
5 Capacity to balance while walking on two feet
6 Brain size and complexity allowing advanced thinking
7 Brain-nervous system control over heart and respiratory actions (e.g. as practised by Yogis)

List B

Capabilities of man

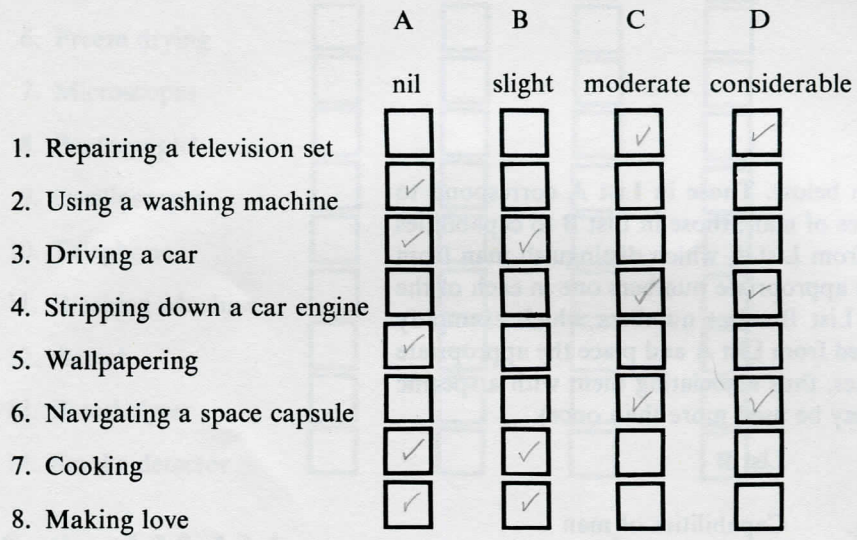
1 Debating
2 Shaving
3 Running
4 Holding food in hands whilst eating
5 Philosophizing
6 Staying under water for short periods
7 Turning metal or wood on a lathe
8 Addressing a public meeting
9 Whistling
10 Telephoning
11 Jumping
12 Communication between individuals by writing
13 Development of mathematics
14 Manufacture and use of tape recorders

2	1, 5, 8, 10, 12	4	2 4 7 12 ~~14~~	6	1, 5, 8, 10, 12, 13, 14

Section 1.4

Question 7 (*Objective* 9)

Indicate the amount of scientific and technical knowledge which you think is necessary for each of the following human skills by placing a tick in one of the boxes opposite each skill.



Sections 1.5, 1.6, 1.7

Question 8 (*Objective* 1)

Each of the numbered definitions listed on the left below may describe one or other of the lettered terms or expressions listed on the right. Put the appropriate letter from the list on the right, in the box provided against each definition. A letter may be used more than once. Place an X in boxes where the definitions do not correspond to any of the terms or expressions.

1 A conclusion which must be accepted because the premisses upon which it is based have been accepted [H]

2 Reasoning from parallel instances [D]

3 An attempt to discover "natural laws" [A]

4 Endowment of a non-human thing with human qualities [C]

5 The application of certain rules of procedure and enquiry accepted as being scientific [A]

6 A generalizing inference from strong circumstantial evidence [G]

7 A belief derived from faith [X]

8 The development of the results of science to aid man in his attempt to control his environment [B]

9 A supposition entertained in order to account for a phenomenon [F]

A Science (several meanings)

B Technology

C Anthropomorphism

D Analogy

E Theory

F Hypothesis

G Induction

H Deduction

Self-Assessment Answers and Comments

Question 1

Answers. 1, G; 2, B; 3, F; 4, C; 5, F.

Comment

These figures, mentioned at the beginning of the first section of the unit, emphasize not only the enormous influence of science and technology in the last century, but also the vast increase in the *rate of change* of scientific development

Question 2

Answers and comments

The following are obvious, *direct* effects of science and technology in the sense that they would never have happened otherwise: 2, 5, 13, 14.

The following are indirect effects, in the sense that they would logically be attributed to other courses, but are in effect in our civilization side-effects (or spin-out) effects of scientific developments:

3—probably mainly due to TV

7—in fact due to ready current availability of central heating systems (but Roman villas had central heating in their fireplaces, and stone water bottles)

9—due possibly to more efficient and rapid communication systems

10—The Open University was originally thought of as the 'University of the Air'—the idea was that the *main* media of teaching would be TV and radio. Clearly such an idea would not have arisen but for these technological developments.

Question 3

You are referred to the examples given with the question.

Question 4

Answers and comments

There is certainly room for argument about some of these, but one possible set of answers might be:

1 None—but you might try to argue for C.

2 Clearly A—this instrument makes distant objects appear nearer and hence more visible.

3 None—but, again you might argue that you cannot separate alcohol from a mixture of alcohol and water by using your bare fingers, but you can with a distillation apparatus—hence, perhaps C.

4 None.

5 In a superficial sense, B. You can *hear* John Arlott commenting on a cricket match at Trent Bridge even though you are in London. So in a way, a radio amplifier extends your sense of hearing.

6 None.

7 Obviously A.

8 You might have had to look this up in the glossary, or in a dictionary. If you did, you will have found that it is an instrument for copying on any scale, for instance a map, diagram or chart. Since it augments the graphic skill of a man's hand (and eye) you might say mainly C with a touch of A.

9 An oscilloscope is a highly versatile instrument, which can display in the form of a visible trace on a screen a wide variety of phenomena or processes, provided some intermediate device is used to turn the primary phenomenon or process into an electrical signal, or impulse, or a series of these. Considered in this way it is an instrument that can 'control' senses A, B and C. But perhaps its most important use is to expand *time*, that is to display a quasi-permanent image of a process that happens in a time too short for it to be observed (seen, heard or felt) directly. None of the senses A to E sense time as such.

10 Superficially B, in a similar sense to 5.

11 None, or perhaps at a pinch C. See comment 3.

12 Obviously B, as you will agree if you look it up in the glossary. It is, or was, used to detect submarines, and locate their approximate position, by sound waves.

13 Presumably D. It is supposed to be a more objective and quantitative device than a policeman's nose.

14 Likewise D. Essentially, one can leave an 'automatic' nose in an unattended building, thereby extending the range of one's sense of smell.

Question 5

Answers and comments

4, 6, 9, all demonstrate a dependence on brain memory; 3, 7 and 10 are reflex acts that do not depend on brain memory. 4, 6 and 9 demonstrate reliance on a *learning process*; 1, 2, 5 and 8 are examples of *innate* behaviour patterns.

Question 6

Answers

2, 4 and 6 in List A distinguish man from all other organisms.

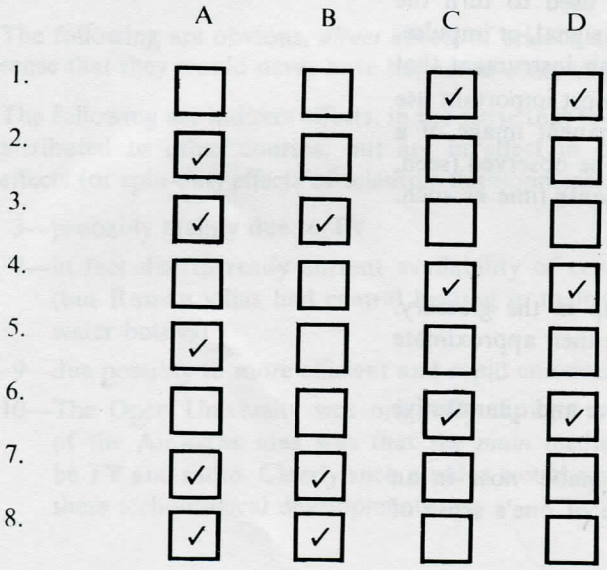
Associated items from List B are:

for 2 in A: 1, 5, 8, 10, 12
for 4 in A: 2, 4, 7, 12, 14
for 6 in A: 1, 5, 8, 10, 12, 13, 14.

Question 7

Answers

One could probably argue about the degree of scientific and technical knowledge required for some of these skills, but perhaps most people would agree with the options suggested below:



Evidently the activities are not specified with sufficient precision to permit a definitive assignment to each one. For instance 4 would require more technical knowledge if you had to do it without the help of any sort of instruction manual (particularly if you were expected to put the engine together again afterwards). Even activity 6 might require only moderate scientific or technical knowledge if the capsule's flight were to be practically entirely controlled by the ground-based control station. This question is itself an example of the fact that to define a *skill* you have to include some statement of the *conditions* under which this task is to be performed.

Question 8

Answers

1, H; 2, D; 3, A; 4, C; 5, A; 6, G; 7, X; 8, B; 9, F.

Comment

1 A deduction is a statement which *necessarily* follows from a proportion or premiss, irrespective of whether or not that proposition or premiss is true.

3 and 5 See page 11 of this Unit.

4 It is very common to find people endowing their pets with human attributes without the slightest justification. A more extreme example of anthropomorphism is the behaviour of a child towards his teddy bear.

6 An induction is a generalization which, although highly probable, can never be proved but only disproved. An induction is, in strict logic, indefensible. But nevertheless most creative thinking uses induction freely.

9 One makes an hypothesis to provoke one to think or experiment further. The trick used by scientists is to make a provocative hypothesis and then to design experiments aimed at disproving it. If the hypothesis is sufficiently important, continued failure by many different experiments to disprove it gradually transforms the hypothesis into a generally accepted theory. Even so, the distinction between an hypothesis and a theory is rather blurred in practice.

Acknowledgements

Grateful acknowledgement is made to the following sources for illustrations used in this Unit:

JONATHAN CAPE LTD. and ALFRED A. KNOPF INC. for two series of drawings showing the development of the brain from gorilla to modern man, from *Man, Time and Fossils*, by Ruth Moore, drawings by Sue Richard. (Fig. 3.)

AEROFLIMS LTD. (Fig. 2a.)
ALDINE PUBLISHING CO. (Fig. 3.)
CAMERA PRESS. (Fig. 2b.)
POPPERFOTO LIBRARY. (Figs. 5a, b and c.)
ZOOLOGICAL SOCIETY OF LONDON. (Figs. 4a and b.)

The Open University

Science Foundation Course Unit 2

OBSERVATION AND MEASUREMENT

Prepared by the Science Foundation Course Team

THE OPEN UNIVERSITY PRESS

Contents

Table A	**List of Scientific Terms, Concepts and Principles**	4
	Conceptual Diagram	6
	Objectives	7
	Introduction	9
2.1.1	**Sounds man cannot hear; light he cannot see**	11
2.2.1	**The nature of sound**	11
2.2.2	The sense of hearing	12
2.3.1	**The nature of electromagnetic fields and light**	15
2.3.2	The sense of sight extended	19
2.4.1	**The need for quantitative measurement**	23
2.4.2	Systems of units	25
2.5.1	**The need to extend the scale of measurement**	30
2.5.2	Measurement of large distances	30
2.5.3	Measurement of small distances	32
2.5.4	Measurement of long time intervals	39
2.5.5	Measurement of short time intervals	44
	Appendix 1 (White) A short description of the principles underlying the photographic technique	47
	Appendix 2 (White) The principle of the microscope	49
	Appendix 3 (Red) Fields	52
	Appendix 4 (Red) Lenses	53
	Appendix 5 (Red) The optics of the human eye	55
	Appendix 6 (Black) Exponential functions	57
	Appendix 7 (White) Glossary	59
	Self-assessment questions	61
	Self-assessment answers and comments	66

Table A

A List of Scientific Terms, Concepts and Principles used in Unit 2*

Taken as pre-requisites | **Introduced in this Unit**

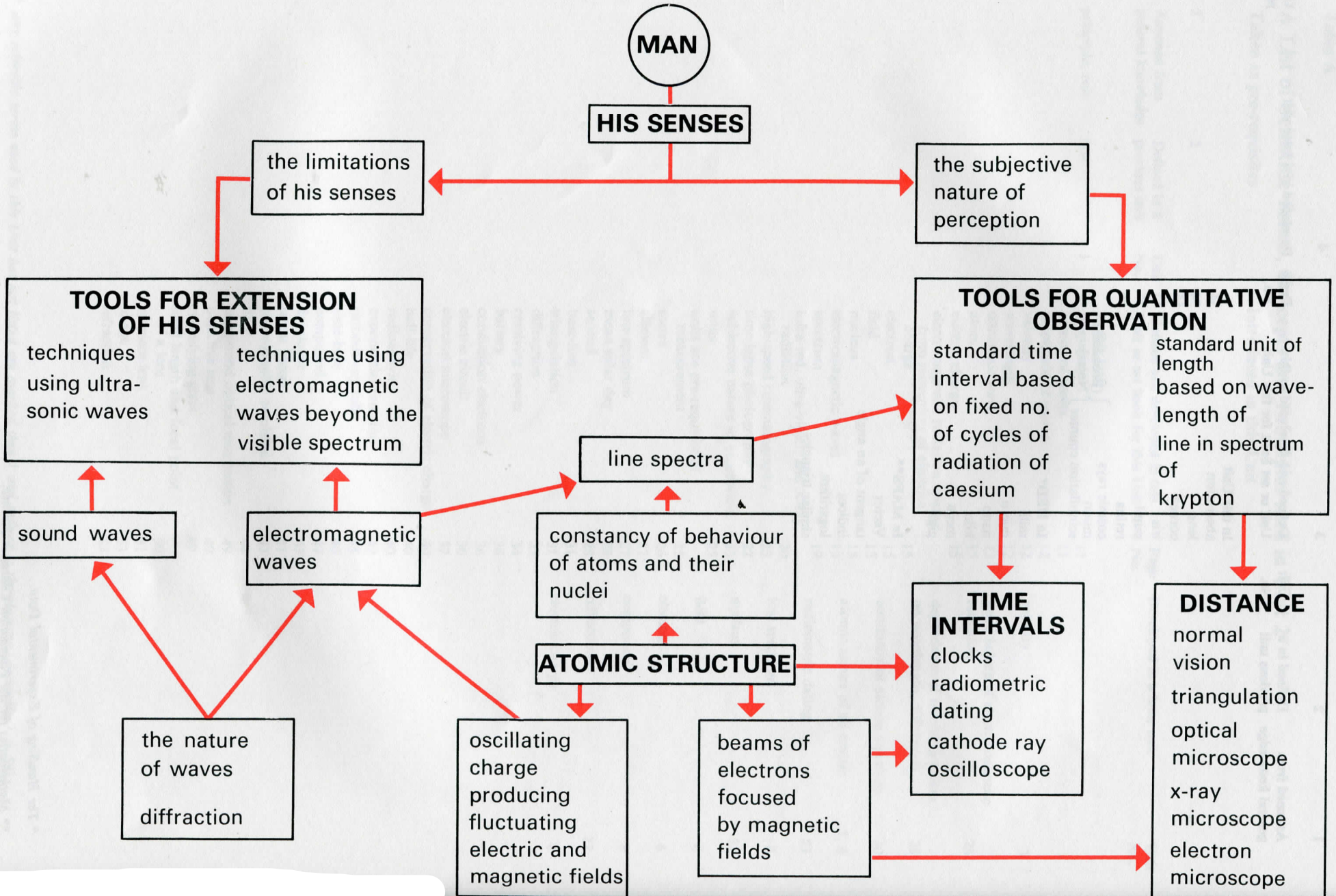
1	2		3		4	
Assumed from general knowledge	**Defined in a previous unit**	**Unit No.**	**Defined and developed in this Unit or set book for this Unit**	**Page No.**	**Developed in a later unit**	**Unit No.**
			In this Unit			
principle, rule	law	1	frequency	11		
			source of waves	11		
			amplitude of waves	12		
			velocity	12	velocity	3
			wavelength	12		
			ultrasonic wave	13	wave nature of electromagnetic radiation	28
			atom	15		
			electric charge	15		
			electromagnetic radiation arising from movement of electrical charge	15	dependence of resolving power on wavelength, diffraction	28
			electron	15		
			field	15	constancy of atomic structure	30
			nucleus	15		
			electromagnetic waves	19	atomic nature of all matter	5, 6
			spectrum	19		
			infra-red, ultra-violet, and X-ray radiation	20	radiometric dating	23
			high-speed cinematography	22		
			time-lapse photography	22	line spectrum	6
			subjective nature of perception	23		
			units	25	nucleus	31
			units are pre-requisite of measurement	25	field	4
			metre	26		
			element	27	element	6
			line spectrum	27		
			mean solar day	28	compound	8
			second	28		
			base-line	31	refraction	22
			triangulation	31		
			diffraction	33	electric charge	4
			resolving power	34		
			battery	36		
			conduction electrons	36		
			electric circuit	36		
			electron microscope	37		
			conservation of electric charge	40		
			half-life	40		
			radioactivity	40		
			exponential function	42		
			cathode-ray tube	45		
			time-base	46		
			compound	47		
			developer	47		
			latent image	47		
			photographic emulsion	47		
			fixer	48		
			compound optical microscope	49		
			objective lens	49		
			magnifying glass	49		
			focal length and focal point of a lens	50		
			eyepiece lens	51		
			image	53		
			refraction	53		

* *Any scientific terms used in this Unit but not listed are marked thus † and defined in the glossary* (*Appendix* 7, *p.* 59).

1	2		3		4	
Assumed from general knowledge	**Defined in a previous unit**	**Unit No.**	**Defined and developed in this Unit or set book for this Unit**	**Page No.**	**Developed in a later unit**	**Unit No.**
			In this Unit			
			aberration	54		
			lens	54		
			cornea	55		
			iris	55		
			pupil	55		
			retina	55		
			cosmic rays, muon, scintillation counter } (Broadcast Notes)			
			In HED*			
			milli			
			micro			
			nano			
			kilo			
			mega			
			giga			
			In MAFS**			
			Vector			
			tangent of an angle			
			indices			
			logarithm			
			similar triangles			

* *The Handling of Experimental Data.*

** *Mathematics for the Foundation Course in Science.*


MAN
HIS SENSES
the limitations of his senses
the subjective nature of perception
TOOLS FOR EXTENSION OF HIS SENSES
techniques using ultra-sonic waves
techniques using electromagnetic waves beyond the visible spectrum
TOOLS FOR QUANTITATIVE OBSERVATION
standard time interval based on fixed no. of cycles of radiation of caesium
standard unit of length based on wave-length of line in spectrum of krypton
line spectra
sound waves
electromagnetic waves
constancy of behaviour of atoms and their nuclei
ATOMIC STRUCTURE
TIME INTERVALS
clocks
radiometric dating
cathode ray oscilloscope
DISTANCE
normal vision
triangulation
optical microscope
x-ray microscope
electron microscope
the nature of waves
diffraction
oscillating charge producing fluctuating electric and magnetic fields
beams of electrons focused by magnetic fields

Objectives

When you have completed the work for this Unit, you should be able to:

1. Define, or recognize adequate definitions of, or distinguish between true and false statements concerning each of the terms, concepts and principles in column 3 of Table A.

2. Given simple unlabelled diagrams, assign names and functional descriptions to the main parts of the following:
   —the human eye
   —a compound optical microscope
   —an electron microscope
   —a cathode-ray oscilloscope.

3. Draw and explain simple ray diagrams to show how the image of an object is focused by:
   —the human eye
   —a compound optical microscope
   —an electron microscope.

4. Give or recognize examples of how man has extended his senses by using:
   —ultrasonic sound waves
   —infra-red rays
   —ultra-violet rays
   —X-rays.

5. State or recognize the approximate ranges of frequency and wavelength over which the following electromagnetic rays extend:
   —radio
   —TV
   —infra-red
   —visible
   —ultra-violet
   —X-ray.

6. Describe how the electric charge within an atom is distributed between its electrons and nucleus. Explain briefly, or recognize, valid descriptions of why two pieces of material rubbed together may become capable of attracting one another.

7. Compare and contrast the mode of generation and propagation of sound waves and electromagnetic waves.

8. Describe the constitution of a photographic emulsion; explain briefly the formation of a latent image, the function of the developer and fixer, and how a final positive image is obtained from the negative image.

9. Name three examples of sensory receptors becoming fatigued by prolonged stimulation and describe the effects produced by such fatigue. Give one reason for believing that some visual illusions arise from the way the brain organizes the information sent to it by the senses.

10. Describe briefly how our system of units evolved and why earlier definitions of the units of length and time were considered unsatisfactory; give the present-day bases of the international standards of length and time.

11. Describe briefly under what conditions electromagnetic fields and waves are produced, and how their existence can be detected. Describe the directional orientation of the field at a point relative to the position and motion of the charge producing it.

12. Describe briefly the behaviour of conduction electrons in a metal when the metal is placed in an electric field, and why a battery is necessary for maintaining an electric field round a circuit. Describe briefly how one can produce a beam of free electrons.

13. State the law of conservation of electric charge.

14. Either by recalling the appropriate formula, or selecting from given formulae, and given relevant data, solve simple problems concerning:

—measurement of distances by triangulation
—the size of an image formed by a lens
—exponential variation
—the relationship between velocity, wavelength and frequency of waves.

Introduction

All living organisms depend to some extent on their ability to observe their surroundings. Man is no exception. His observations must, of course, come to him through his senses, but he has come to realize that his senses are limited, both in their scope and reliability. By the ingenious invention of suitable tools and techniques, however, he is to a large extent overcoming these limitations, and in so doing has opened up a rich new experience of the world in which he lives.

With these tools he has been able:

(a) to extend the range of his senses so as to detect sounds and light that otherwise would not be accessible to him;
(b) to quantify his observations and make them into reproducible measurements;
(c) to investigate phenomena occurring on scales of length and time that are either much greater or much smaller than those that come within his natural domain.

This then briefly is the theme of this Unit. It is a very broad one as you can well imagine, so naturally we have had to be very selective in our material. At first we thought we would take you to the very frontiers of, for example, small scale measurements. This turned out to be none too easy at this early stage of the Course. In any case we thought it would be more appropriate to deal with these techniques later when we discuss elementary particle physics.

The illustrations we eventually decided upon were chosen for a variety of reasons. For example, in the measurement of short time intervals we describe a cathode-ray oscilloscope. This is an instrument very similar in principle to your television set. Relying as we do at the Open University on television for so much of our teaching we felt you ought to know something of what goes on behind that screen. We shall also be making extensive use of photography as a means of communicating with you—both in the form of photographs issued with the correspondence texts and as films shown on television. This was one of the reasons why we chose to describe the photographic technique. Our decision to concentrate on the microscope as a means of measuring small distances was influenced by the fact that you were to be lent a microscope for your own use and so should obviously know how it works.

A wide-ranging topic like measurement also allows us to fill in a great deal of background knowledge. For example, we can cover topics such as the nature of sound, the nature of light, elementary electricity, simple optics, etc.

If you already have a firm grounding in basic physics, then you will find much of this Unit easy going. Don't miss any of it out, though, as we want you to follow through the general theme—but do feel free to take the material at a brisk pace. Providing, of course, you can manage the tests at the end of the Unit, you may then make an early start on Unit 3; this introduces Einstein's Special Theory of Relativity and it almost certainly will give you scope to stretch yourself. If on the other hand your scientific background is not so good, then you have a considerable amount of ground to cover this week. Don't let this deter you though—if you go about it conscientiously, taking careful note of the meaning of each new term as it is introduced, you will get through it perfectly well!

Section 1

2.1.1 Sounds Man cannot Hear, Light he cannot See

You saw in Unit 1 how important is the ability to make observations. Even simple organisms need to have some means of observing the surroundings. If these means were not available it would be impossible to locate food and avoid predators.

In man, this process of observation goes much further. Each individual can draw not only from his own personal experiences but also from that of others. He has access to a large accumulated body of information which is continually growing. This ever deeper insight into the workings of nature in turn leads to an increasing ability on man's part to manipulate the environment.

But much of the information he wishes to acquire from his surroundings is not directly accessible to his senses. There are sounds his ears cannot hear; there are forms of light his eyes cannot see.

We have already mentioned how he has developed ways of overcoming some of the limitations of his body; he fashions tools that develop his manipulative skills and others that increase his physical strength. We now consider others that add to his powers of observation—they are extensions to his senses.

Section 2

2.2.1 The Nature of Sound

We begin with the sense of hearing. Hearing probably comes second only to sight in its importance as a means of observing the environment. It centres upon the ear's ability to detect sound waves, so we first describe the nature of sound.

Sound waves are generated by a moving body called a *source*. As it moves, it pushes against the surrounding medium (usually the air) and in so doing causes a disturbance. We can take as an example of a source a vibrating metal strip clamped at one end, as in Figure 1. Its to-and-fro motion sets the nearby layer of air particles moving.

What do you expect the particles in the nearby layer to do now:
(a) move right away from the source, passing through the rest of the medium;
or
(b) transfer the disturbance to the next layer?

They in their turn push against the next layer and rebound and so on. Like the metal strip itself, the individual air particles oscillate about their mean† position (the position they would occupy if the metal strip were not moving), while the disturbances initiated by the source progressively move outwards. The disturbances consist of regions where the *density* (the number of particles in a given volume) is successively greater and less than the mean density. A series of repeated disturbances is called a wave train. Figure 2 shows the positions of these disturbances at various times after the strip has been set vibrating.

The number of high density regions passing any given point in a given time is called the *frequency* of the wave train and we denote it by *f*. It has the same value as the number of complete to-and-fro motions of the source in the given time.

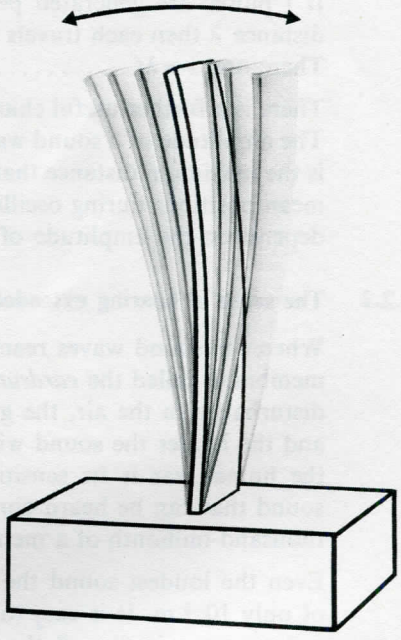


Figure 1

A source of sound—a vibrating metal strip.

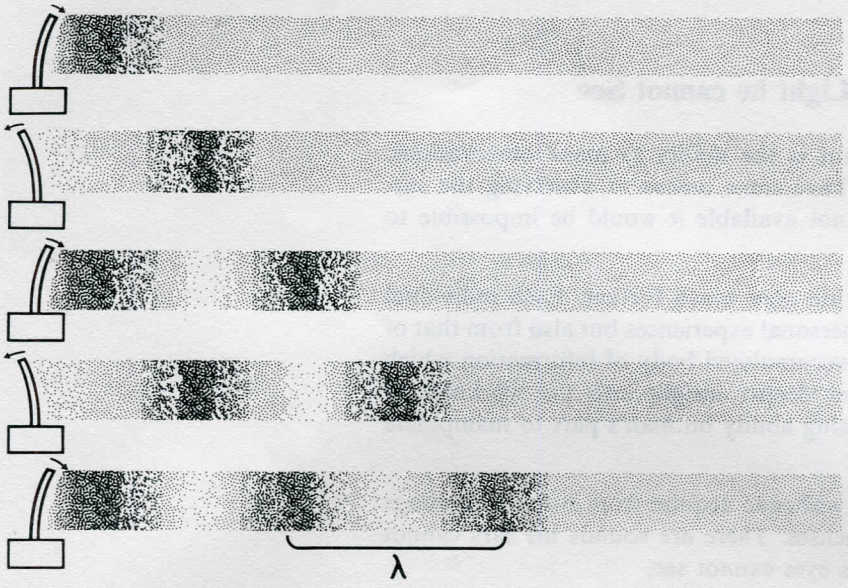


Figure 2

Diagram showing successive regions of high and low density being generated by a vibrating metal strip. The wavelength, λ, is illustrated on the final drawing.

If the vibration of the source is maintained at a steady rate, then it is clear that the distance between successive pulses of high density will always be the same. This constant distance is called the *wavelength* (Fig. 2) and is denoted by λ (a Greek letter pronounced 'lam-da'). It also represents the distance between successive regions of low density. The wavelength obviously depends upon the *velocity*, v, with which the pulses are moving away from the source (i.e. the distance they travel in a given time).

What relationship do you expect between f, v, and λ?

If f pulses are generated per second and they are each separated by a distance λ then each travels a distance λf per second.
Therefore, $v = \lambda f$ ..... (1)

There is a further useful characteristic of a wave, and that is its *amplitude*. The amplitude of a sound wave is related to the loudness of the sound and is the maximum distance that the particles of the medium move from their mean positions during oscillation. In the case we are considering this will depend on the amplitude of the motion of the metal strip.

2.2.2 The sense of hearing extended

When the sound waves reach the ear, they set vibrating a thin stretched membrane called the *eardrum*. Naturally the greater the amplitude of the disturbance in the air, the greater will be the movement of the eardrum and the louder the sound will appear. One of the remarkable features of the human ear is its sensitivity to very small amplitudes. The faintest sound that can be heard corresponds to a displacement of one hundred-thousand-millionth of a metre* (i.e. 10^{-11} m)**.

Even the loudest sound the ear can tolerate has a maximum amplitude of only 10^{-5} m. It is easy to appreciate from these figures that the extraordinary sensitivity of the ear makes it exceedingly difficult for the designers of aircraft, for example, to keep the extraneous noise down to what might be considered a tolerable level.

The ear, of course, not only differentiates between faint and loud sounds, but also between sounds of different frequency. If the pulses arrive at the rate of about 20 cycles† per second or less, the ear is able to distinguish the separate pulses. Above this frequency the effect produced is that of a continuous note—the higher the frequency, the higher the note.

* 1 *metre* = 39.4 *inches*.

** *The amplitude for the faintest audible sound depends upon the frequency involved. The figure quoted is for a frequency of* 1 000 *cycles per second.*

It is at this stage that one encounters a limitation of the ear. Above about 20 000 cycles per second, sounds can no longer be heard. The actual value of the frequency at which the cut-off occurs depends upon the individual and varies with age. Waves with frequencies greater than 20 000 cycles per second are classified as *ultrasonic waves*.

The ears of some animals respond to much higher frequencies than 20 000 cycles per second. This is why it is possible to have special high frequency whistles that produce waves audible to dogs, but not to humans. Thus man finds himself up against a barrier—a barrier imposed by a limitation of one of his senses. He has even to face the fact that some of the so-called lesser animals have a natural ability that excels his own in this respect! If he were to extend the range of his hearing to higher frequencies with an instrument of some sort, would it be to his advantage? One has only to consider the bat to find an answer (Fig. 3).

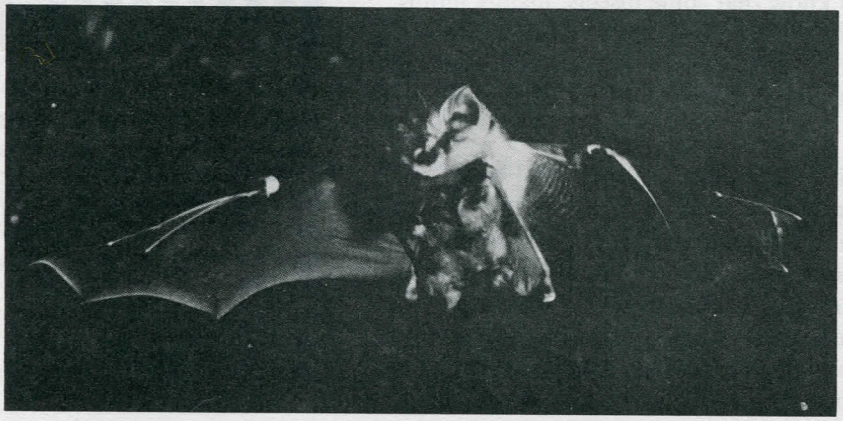


Figure 3

A bat in flight.

Bats have a remarkable ability to fly at speed in and out of trees in darkness. As they fly they generate ultrasonic waves. How the waves are generated is not fully understood, but it is believed that they radiate from the bat's nose during both inhalation and expiration. Thus the sound is emitted during the complete breathing cycle. The frequency generated is about 50 000 cycles per second and this corresponds to a wavelength of 0.7 cm. Sometimes the frequency may be as high as 150 000 cycles per second.

The bat issues short bursts of these ultrasonic waves in a well-defined forward direction. In normal cruising flight, the pulses are emitted at a rate of 5 to 20 bursts per second with silence in between. These pulses bounce back to the bat from any objects in its path (Fig. 4). Knowing the direction from which the echo comes, and deducing the distance from the time that has elapsed between the emission of the original burst and the arrival of the echo, the bat can locate the object.

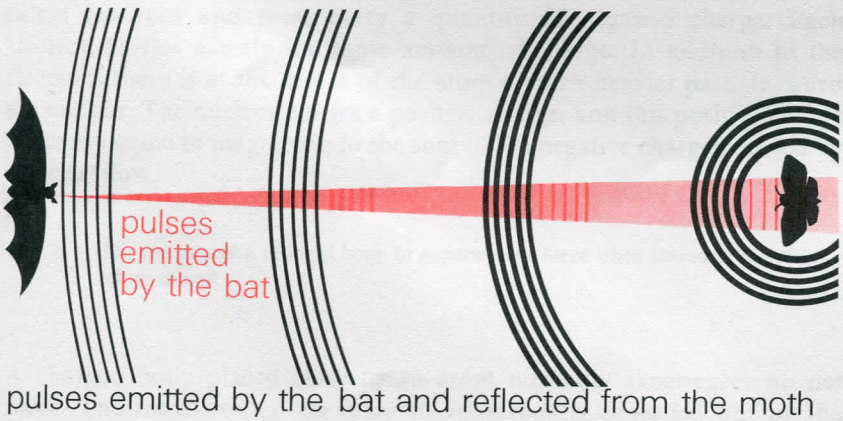


Figure 4

Diagram illustrating the way bats emit short pulses of ultrasonic waves in a narrow forward cone. These are then reflected back to it from objects in its path.

But why use ultrasonics—why not ordinary sound? The reason is to be found in Figure 5. Here water waves are seen striking objects of various sizes. You can see that when the object has dimensions almost as small as the wavelength, the wave train is hardly disturbed at all. It is only when the size of the object is greater than a wavelength that it reflects an appreciable amount of the wave.

Waves of all kinds behave in similar ways. What is true of water waves is also true of sound. The size of the object to be detected with the sound waves sets the limit on the size of the wavelength. The bat uses a wavelength that is ideal for detecting branches, twigs, and leaves which can then be avoided, and insects in flight on which to prey. The wavelength is small enough to get good reflection but at the same time is not too small. Ultrasonics of shorter wavelength dissipate their energy and die out more rapidly, so there is an advantage in not having the wavelength any shorter than necessary. (You might like to know that some insects, mainly moths, have developed ears that are sensitive to the frequencies emitted by bats, and so they receive a warning of the bat's approach. Some moths are even capable of producing their own ultrasonic waves and these appear to deter the bat from coming too close.)

Man has now succeeded in copying very effectively the bats' technique. We can generate and detect ultrasonic waves, for example, by making crystals vibrate rapidly. Because of their dimensions and physical properties, these crystals have a natural tendency to vibrate much more quickly than any metal strip or eardrum. They provide the means of both generating and detecting ultrasonic waves, and can be combined into a system similar to that used by a bat. Such man-made systems are used at sea, for example, and help in the location of submarines and shoals of fish. They even distinguish between shoals of large and small fish. Systems are being developed for the guidance of the blind and these have now reached the stage where they can successfully be used to locate pencils at a distance of several feet.

Ultrasonic waves can also probe into objects that are opaque to light, such as metal castings. The waves can be passed through the object and detected on the far side. For normal castings there will be some characteristic amplitude for the transmitted wave. If however a specimen has any flaws or impurities in it, the waves will be partially reflected from them. This leads to a detectable reduction in the amplitude of the transmitted wave. An abnormally weak, transmitted wave is thus an indication that a specimen is faulty.

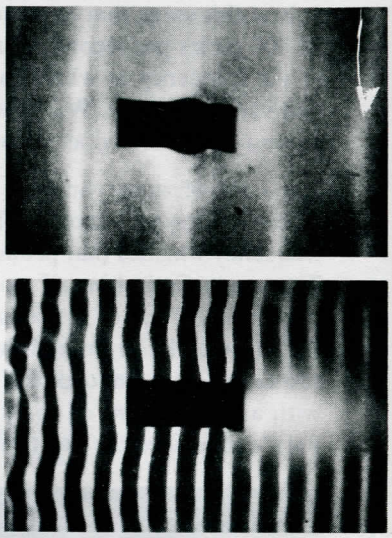


Figure 5

Water waves of different wavelengths advancing from the left and meeting an object. The smaller the wavelength the greater the disturbance caused to the wave train. Note there is an almost total absence of shadow in the upper photograph.

2.3.1 The Nature of Electromagnetic Fields and Light

Impressive as these new skills may be, they compare hardly at all with the enormous developments that have stemmed from man's extension of that most important of all senses—sight.

The eye, like the ear, responds to the reception of a wave. But this wave is quite unlike that of sound. Sound needs a medium; without the air particles, or particles of some other gas, solid or liquid, there would be nothing to carry the vibrations from the source to the detector.

Light waves, however, need no medium. They travel across empty space—if they did not one would never see the Sun or stars. So what do they consist of? To answer this we need to introduce some basic ideas about electricity and magnetism.

You are no doubt familiar with the behaviour of magnets. If a small steel pin is placed close to a magnet, such as the one provided in your Home Experimental Kit, it is pulled towards it, i.e. it experiences a force of attraction.

Do you expect the force of attraction to depend upon the distance between the pin and the magnet?

The closer the pin to the magnet the greater is the force. The magnet is, therefore, exerting some kind of influence that pervades space, the strength of this influence falling off with distance. Thus, corresponding to each position in space relative to the magnet, there is a characteristic force that will be exerted on the pin were the pin placed in that position.

The same kind of thing is found with electricity. Electrical effects are produced by what are called *electric charges*. There are two types. Bodies carrying charges of the same type repel each other whereas bodies with dissimilar charges attract each other. One type of charge is called 'positive', the other 'negative'. The choice of name is quite arbitrary.

Two positive charges repel/attract?
Two negative charges repel/attract?
A positive charge repels/attracts a negative?

Two positive charges repel each other as do two negative charges, and a positive and negative charge attract each other.

All matter is composed of *atoms*. Each atom, as you will see in detail later, has a composite structure. Some of its component parts are tiny particles called *electrons* and these carry a quantity of negative charge. Each electron carries exactly the same amount of charge. In addition to the electrons there is at the centre of the atom a much heavier particle called the *nucleus*. The nucleus carries a positive charge, and this positive charge is exactly equal in magnitude to the sum of the negative charges carried by the electrons.

Do you expect a charged body to experience a force when placed in the vicinity of an atom?

A charged body placed close to an atom normally experiences no net force. The force exerted on it by the electrons is cancelled out by the

equal force exerted by the nucleus in the opposite direction. The atom is therefore regarded as having a net charge of zero—it is said to be 'neutral'.

Atoms are not 'indivisible', as was once thought. Depending upon the type of atom involved, it is more or less easy to remove an electron from it. So, if two dissimilar materials are vigorously rubbed against each other, the one made of atoms with the more loosely bound electrons is likely to lose some of them. They are transferred to the other piece of material. In this way both objects acquire an unbalanced charge—one by virtue of having too many electrons, the other through having too few.

It is then that the electric charges work to produce large-scale effects that are easily visible. These include the way in which strands of hair sometimes repel each other and stand apart when a comb has been passed through them. A lightning flash is another example; this is an electrical discharge initiated by accumulated charges in the atmosphere.

Thus, an electrically charged object exerts an influence throughout space, in much the same way as a magnet does. Once again it is found, when the phenomenon is investigated quantitatively, that the force exerted on another body depends upon the distance to that body, and it progressively diminishes the greater the distance.

In this way, one comes to think in terms of *fields**—electric and magnetic (and later we shall add gravitational and nuclear). An electric charge is thought of as being surrounded by an electric field. This is a 'condition' produced by the charge and this condition pervades the surrounding space. It can be assigned a magnitude and direction at every point. The electric field at a point is specified (a) by the strength of the force that a body carrying a given amount of positive charge would experience if placed at that point, and (b) by the direction in which that force would act. If the magnitude of the force on the body when at one point is twice that when placed at some other point, then the field at the first point is said to be twice that at the second.

This term 'field' is a difficult one to explain satisfactorily. Nonetheless the concept is a most useful one for it allows the analysis of force problems to be divided into two convenient parts: In the first place, one charged body, A, is regarded as being the *source* of the field—the field is there whether or not there are any other bodies to be affected by it—while any other body, B, is thought of as being *acted upon* by the field. The force on B arising from the action of this field can then be studied. Similarly the force on A due to the presence of B can be studied by considering B to be the source of a second field—one that acts on A.

Clearly, while an electric charge remains stationary, the direction and magnitude of the field it produces remains constant at all points.

What do you expect will happen to the field if the charged body is moved to another position?

If, however, the charge is moved to another position the field has to change to one that is appropriately centred around the new position. Suppose a charge is initially situated at position A in Figure 6. The field at the equally spaced points a, b, c, d . . . acts in the directions indicated by the solid arrows, i.e. directly away from A. If the charge is moved to position B, the field has to be directed away from this new position, i.e. acts along the dotted arrows.

* *We are using the word 'field' here in a very special way. Forget other connotations the word has for you, such as ones closely associated with the idea of area, field of view, grass, etc.*

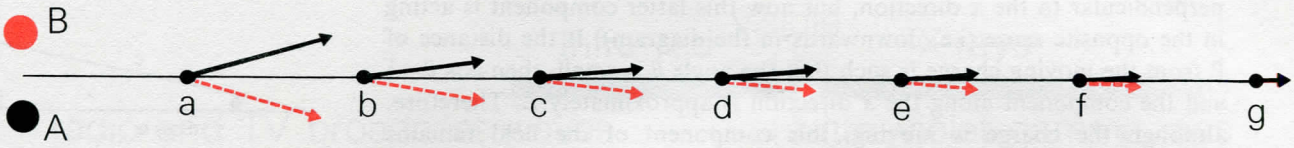


Figure 6

When a charge is situated at position A, the electric field at points a, b, c, d, e, f, g, . . . *act along the directions indicated by the solid arrows. When the charge is at position B the field acts in the direction of the dotted arrows. We show the length of the arrows becoming progressively smaller to remind you that the strength of the field falls off with distance. (The rate at which the field strength falls off with distance is in fact much more rapid than is indicated by the lengths of the arrows.)*

However, this reorientation of the field does not take place instantaneously. It is found that a certain time, t, has to elapse before the field at point, a, changes. If the distance of the charge to point a is x, and that to point b is $2x$, then it takes time $2t$ before the appropriate change is initiated at point b; $3t$ for point c, etc. In other words the disturbance to the field moves out from the moving charge with a certain velocity given by x/t.

Having moved the charge to the second position, it could immediately be returned to its original position so as to restore the *status quo*.

The charge is back to its original position, but is everything restored to its original state?
Was that last sentence correct?

However, the *status quo* has not been exactly restored. This is because the disturbance produced in the field still carries on out into space. For example in Figure 6, having moved the charge to position B and allowed the disturbance to reach point d, one could return the charge to position A. The initial disturbance to the field is unaltered by this latest movement of the charge and still carries on from d to points e, f, g . . . (with, of course, a second disturbance now following it some way behind corresponding to the movement of the charge from B back to A). Here indeed is a vivid example of the usefulness of the field concept. By only looking at the charge after the movement is completed, one cannot tell that anything has changed. Indeed, as far as the charge is concerned, nothing *has* changed—it is back where it started. A study of the charge alone therefore is inadequate to explain why some time later a second charge at, for example, point g suddenly experiences an abnormal force. Once having introduced the concept of a field that is distinct from either of the two charges, however, one can explain the behaviour of the second charge in terms of a disturbance to the field.

Now, instead of keeping the source stationary after bringing it back to its original position, one can, if so desired, repeat this movement over and over again at a steady rate. Suppose in Figure 7 the charge oscillates between positions A and B. What is the effect of this on the field at P?

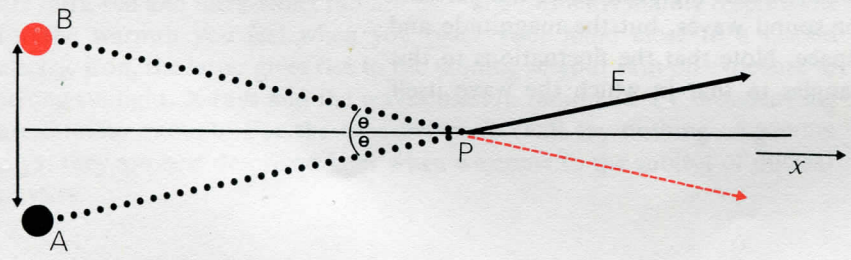


Figure 7

The field at P appropriate to the charge being in position A will have a strength, E, directed along the solid arrow. This can be regarded as being made up of two components $E\cos\theta$ along the x direction and $E\sin\theta$ at right-angles to x (Fig. 8).* When the charge is in position B, the field has the same strength, E, but is directed along the dotted arrow (Fig. 7). This can be resolved once again into $E\cos\theta$ along the x direction and $E\sin\theta$

* *Refer if necessary to* MAFS, *section* 4.D.7.

perpendicular to the x direction, but now this latter component is acting in the opposite sense (i.e. downwards in the diagram). If the distance of P from the moving charge is such that the angle θ is small, then $\cos\theta \approx 1$ and the component along the x direction is approximately E. Therefore, although the charge is moving, this component of the field remains essentially constant. The interesting component is the one perpendicular to x; this is the one relevant to our discussion on the nature of light. It fluctuates in value. (Actually its magnitude exceeds the value $E \sin\theta$ expected on the basis of Figures 7 and 8. This is because the *motion* of the particle generates an additional field in this direction.) Figure 9 shows a graph of how the strength of this field changes with time at a given position.

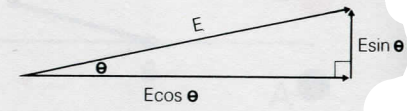


Figure 8

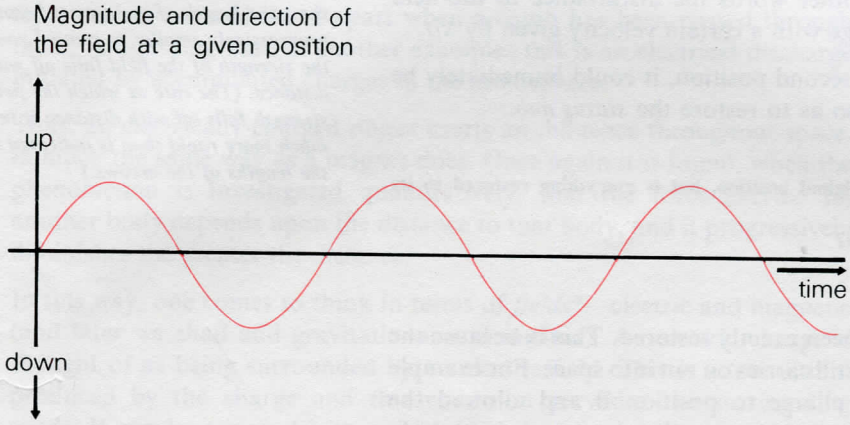


Figure 9

Diagram showing how the fluctuating part of the field at a given point on the x axis *varies with time when the charge oscillates between positions A and B* (see *Figure* 7).

How could we check experimentally that this was the way the field varied?

This can be demonstrated experimentally by placing a small charged body at point P and measuring how the force on it varies with time. If it were free to move at right-angles to the x direction, it would oscillate up and down. In fact it would behave rather like a cork bobbing up and down on the surface of some water over which a wave was passing. Indeed, can this fluctuating part of the field be regarded as a wave of some sort?

To answer this, imagine there to be a number of small charged bodies strung out all along the x direction of Figure 7. At a given instant of time, the forces exerted on each of these bodies by the field are measured. These measurements in turn give the values of the fluctuating field at each position. Figure 10 shows the sort of result one might get.

So in the way the electric field varies, both with time at a given position and with position at a given instant of time, it behaves like a wave train. The quantity that is altering is not of course the position of the particles of a medium, as was the case for sound waves, but the magnitude and direction of the electric field in space. Note that the fluctuations to this field are in a direction at right-angles to that in which the wave itself is moving (the x direction).

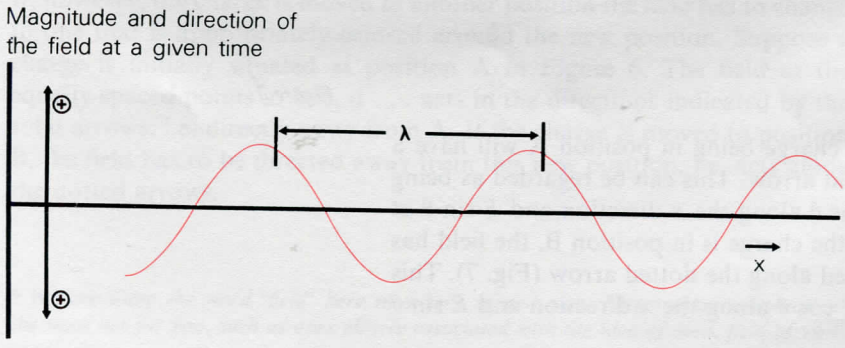


Figure 10

Diagram showing how at a given time *the fluctuating part of the field varies from point to point along the x axis when the charge oscillates between positions A and B* (see *Figure* 7).

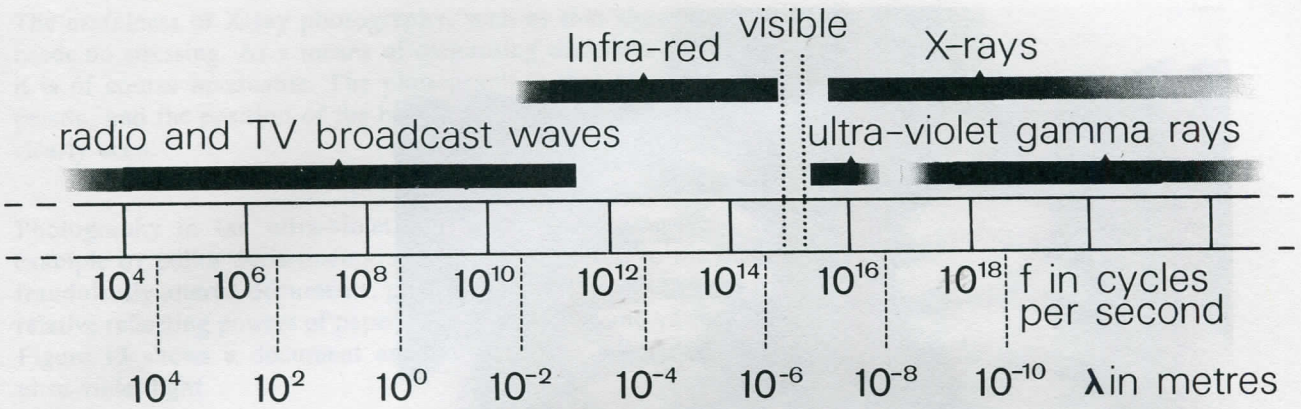

Infra-red
visible
X-rays
radio and TV broadcast waves
ultra-violet gamma rays
10^4 10^6 10^8 10^{10} 10^{12} 10^{14} 10^{16} 10^{18} f in cycles per second
10^4 10^2 10^0 10^{-2} 10^{-4} 10^{-6} 10^{-8} 10^{-10} λ in metres

Figure 11

The electromagnetic spectrum.

The number of times per second that the fluctuating component is found to have a maximum value in the same direction at a given point is the same as the number of to-and-fro vibrations of electric charge; it is the frequency of the wave. The wavelength of the wave shown in Figure 10 is then related to this frequency and to the velocity of the wave by the same formula as applied to the sound waves.

Do you remember the formula?
If not turn back and look it up.

If you are still not absolutely happy about this concept of fields, turn to Appendix 3 (Red) (p. 52).

You might find the analogy described there helpful.

As you will see in Unit 4 there is a very intimate relationship between electric and magnetic fields. In fact, whenever an electric charge is moved, a magnetic field is produced. The oscillating charge in Figure 7, therefore, not only produces fluctuations in the electric field but also gives rise to a fluctuating magnetic field. These magnetic disturbances travel out with the electrical ones and with the same velocity and frequency.

Together, the electric and magnetic disturbances are called *electromagnetic waves*. It is these waves, or at least a restricted class of them, that the eye detects. There is no theoretical limitation on the frequencies allowed to electromagnetic waves and the phrase 'electromagnetic spectrum' has been coined to embrace all vibrations of this type. The word 'spectrum' implies that one is distinguishing between various types of wave according to their frequency (or wavelength).

Figure 11 displays the electromagnetic spectrum. As can be seen, it includes a whole variety of phenomena. Just beyond the visible region there are infra-red and ultra-violet radiations.* The former is mainly responsible for the warmth you feel when you place your hand close to a heated electric iron, the latter gives rise to the tanning of your skin on exposure to strong sunlight. X-rays and the waves used in radio and TV broadcasting are familiar enough. For the time being we shall say nothing of gamma rays; they are best described later when we come to the subject of nuclear physics.

2.3.2 The sense of sight extended

The eye responds to only a narrow band of wavelengths. Within this range, light of differing wavelength can be distinguished by colour. There will be a colour plate showing the visible spectrum in Unit 6. It extends from

** The word 'radiation' is sometimes thought of as having special reference to 'atomic radiation'. We, however, shall be using it in the quite general sense of 'something emanating or diverging'.*

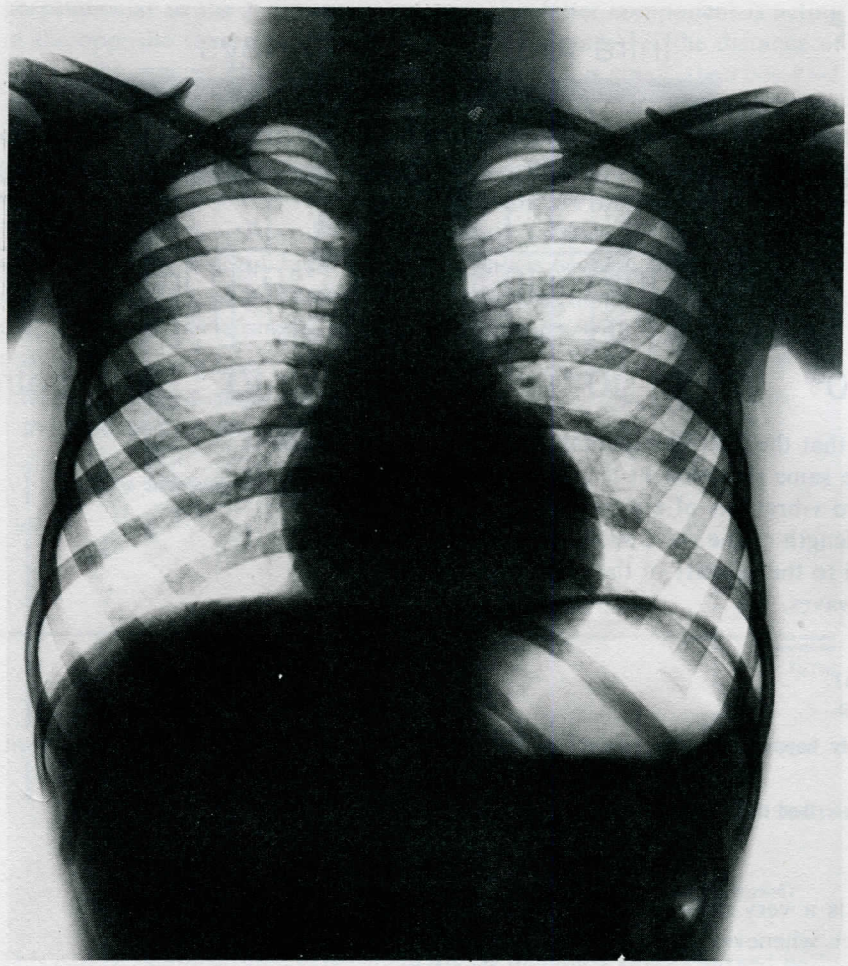


Figure 12

X-ray photograph of a normal healthy person.

violet at one end, corresponding to a wavelength of 400×10^{-9} m, to red at the other, with a wavelength of 700×10^{-9} m. Ordinary white light is made up of a mixture of all these colours. Though the eye can differentiate between light of different wavelengths when they are shown separately, it is not a good analyser of mixtures of light of various wavelengths; it cannot 'see' the component colours of white light, only a single impression of 'whiteness'. In this respect the ear is superior; though an entire orchestra is playing the ear, as you know, can analyse the sound and pick out individual notes. A glance at Figure 11 is sufficient to appreciate how much is missed through the human eye having access to only the visible range.* To the scientist much of the most interesting information about his world is contained in electromagnetic radiation outside the range available to his unaided eye. Thus once again he has to overcome a physical limitation of his body by using auxiliary devices.

There are many such devices, each covering a certain spectrum of wavelengths. One of them has a particularly wide range of application—the photograph.

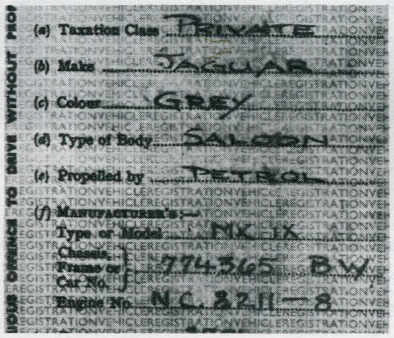
Turn to Appendix 1 for a short description of the principles underlying the photographic technique (p. 47).
Note that this appendix is classed as white-page material and should therefore be read by all students.

The photographic technique can be used to detect radiation of the very smallest wavelengths, right through the visible region and into the infra-red.

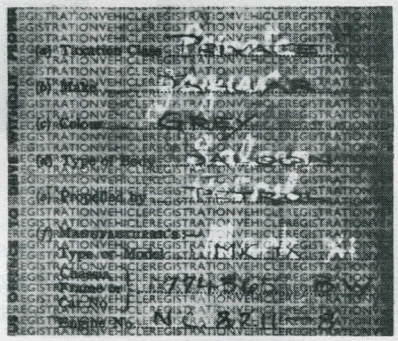
** The visible range shown is that appropriate to the human eye and does not necessarily apply to other animals. Bees and moths, for example, can see in the ultra-violet.*

The usefulness of X-ray photography, such as that shown in Figure 12, needs no stressing. As a means of diagnosing disorders within the body, it is of course invaluable. The photograph is that of a normal healthy person, and the position of the heart on the left-hand side of the body is clearly seen.

Photography in the ultra-violet range has been extensively used, for example by police departments, museums and banks, in the detection of fraudulently altered documents, paintings, etc. This use depends upon the relative reflecting powers of paper, inks, erased inks and various pigments. Figure 13 shows a document as seen under (a) ordinary light and (b) ultra-violet light.


(a) Taxation Class PRIVATE
(b) Make JAGUAR
(c) Colour GREY
(d) Type of Body SALOON
(e) Propelled by PETROL
(f) Manufacturer's Type or Model MK IX
Chassis No. 774365 BW
Frame or Car No.
Engine No. NC. 8211-8

(a)

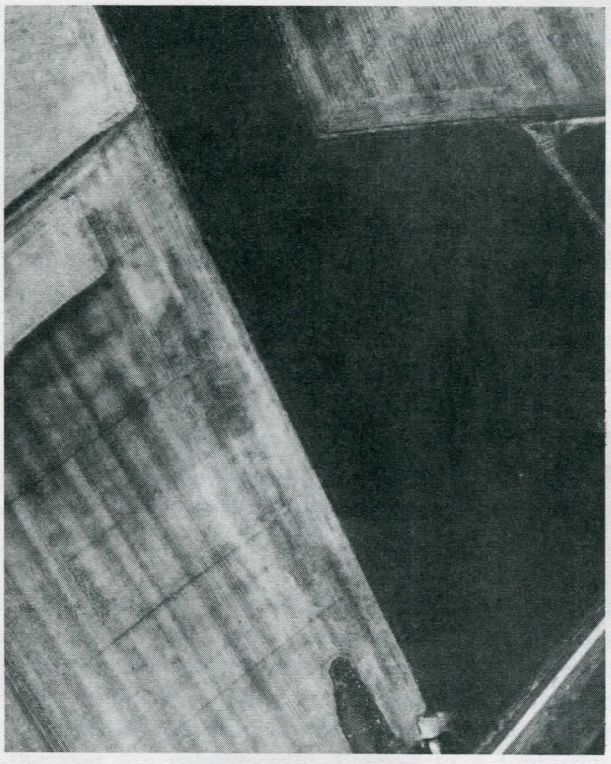


(b)

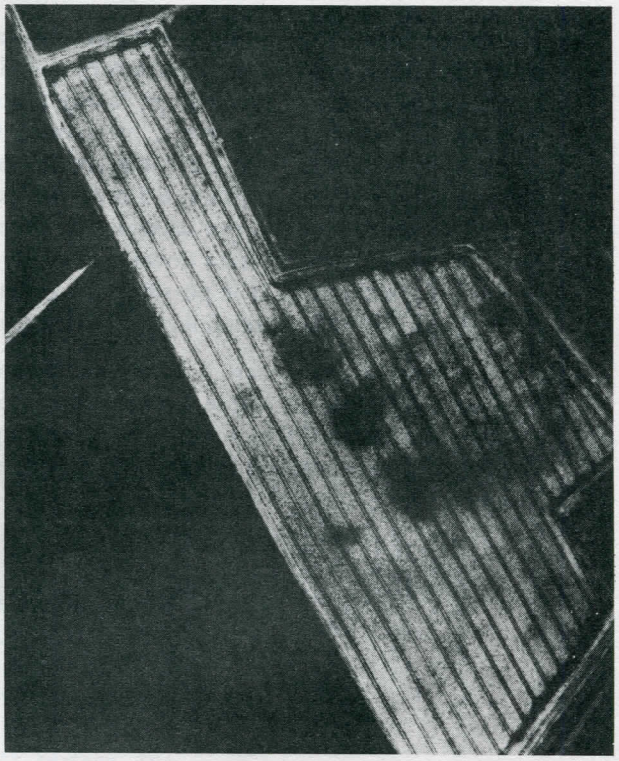
Figure 13

A car log-book viewed in (a) ordinary light, (b) ultra-violet light.

On the longer wavelength side of the visible region photographic methods can extend to the infra-red region as far as about 1.0×10^{-6} m. There are many applications for this type of photography. The aerial pictures of Figure 14 show a potato field in the Cambridgeshire fens as seen (a) by ordinary light and (b) by infra-red light. You can see that whereas the appearance of the field is uniform in ordinary light, there is a very marked



(a)



(b)

Figure 14

A potato field in the Cambridgeshire fens. The dark patches on the infra-red photograph on the right reveal the presence of potato blight not visible on the panchromatic film on the left.

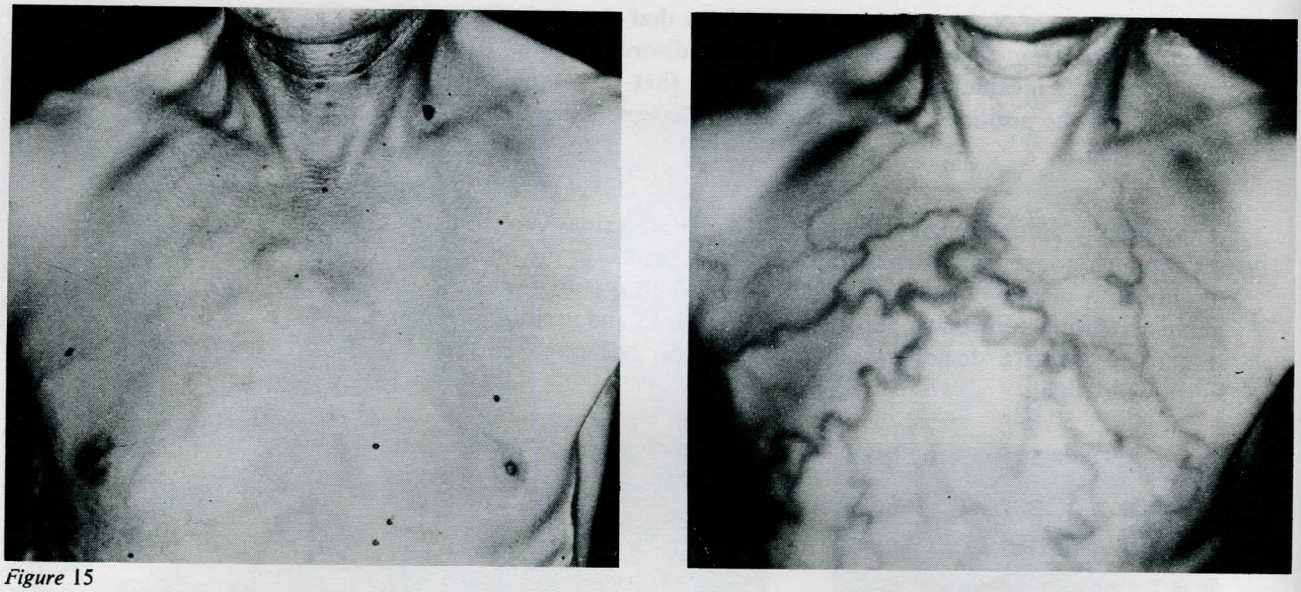


Figure 15

Photographs of a man's chest taken in (a) ordinary light and (b) infra-red light.

difference when it is looked at in the infra-red. The dark patches on the latter reveal the presence of potato blight. In this way aerial photography enables one rapidly to diagnose some diseases in various types of vegetation.

In the medical field infra-red photography can be used in the examination of veins (Fig. 15). Another application is to situations where it is important that the subject remains in 'darkness', for example when one wishes to study the night behaviour of warmblooded animals (or indeed of burglars).

Quite apart from this ability to extend the sense of sight into wavelength regions not accessible to the eye, the photographic technique can be of great use in other respects. For example, the emulsion can be exposed for long periods and thus be made to reveal very faint sources, such as distant stars, which are invisible to the unaided eye.

The development of cinematography extends man's ability to see in yet other directions. If a process occurs very quickly, the eye has difficulty in resolving the details of the changes taking place. The spot of light flying over the TV screen, for example, produces a flicker that is only barely perceptible. This limitation can be overcome with the use of high speed photography. With this technique, frames are exposed at a rate sometimes as high as several million per second. They are then shown at a reduced rate, thus slowing the process down sufficiently for the eye to see what is happening. In this way it is possible to see details of a whole variety of phenomena that otherwise could not be discerned.

At the other extreme, there is time-lapse photography. Here one allows long time intervals to elapse between the exposure of each frame. On showing the film at the usual rate, the process appears greatly speeded up. Glaciers, for example, can be shown to flow like rivers, and plants to grow visibly.

2.4.1 The Need for Quantitative Measurement

You have now seen how man has been able to extend his senses. Unfortunately, his senses are easily misled; they do not always provide reliable information. His sensory receptors can become fatigued or 'adapted' by prolonged stimulation. Here we remind you of a few well-known examples.

You are no doubt already familiar with visual after-images. If you stare fixedly at the cross in Figure 16 for about a minute, and then transfer your gaze to a plain white background, a bright after-image of the cross will be seen. Try it.

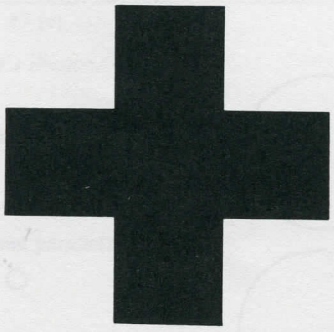


Figure 16

If you carry a heavy suitcase for a long period, on putting it down your arm suddenly feels light.

The sense of taste can be conditioned. The sweetness of a cup of coffee is judged differently depending upon what one is eating at the time—a piece of cake perhaps or something savoury.

We leave it to you to decide whether this is an opportune moment to try this one!

The impression of time 'dragging' or 'flying' depends upon the activity in which one is engaged. Prolonged driving at 70 mph on a motorway can alter one's view of the reasonableness of a subsequent 30 mph speed limit in a built-up area.

Judgement of temperature is subjective. This can be demonstrated by holding your left hand in cold water for a few minutes and your right hand in hot, and then plunging both simultaneously into a bowl of lukewarm water. Try it.

What conclusions do you draw about the hotness or coldness of the water in the bowl based on the reaction of each hand?

Some misleading impressions are due to modifications made to the behaviour of the sense receptors. In response to identical stimulation, receptors transmit different nerve impulses depending upon the level of prior stimulation. Thus, in the case of the after-image effect, prolonged exposure to light reduces the sensitivity of the receptors in the eyes. When the eyes are subsequently directed to a uniformly bright surface, the receptors previously exposed to light transmit fewer nerve pulses than those that were not similarly exposed. The way in which the nervous system operates will be dealt with more fully in Unit 18.

However, not all observational distortions can be attributed to such changes. For example, you will already be familiar with optical illusions such as those shown in Figure 17.

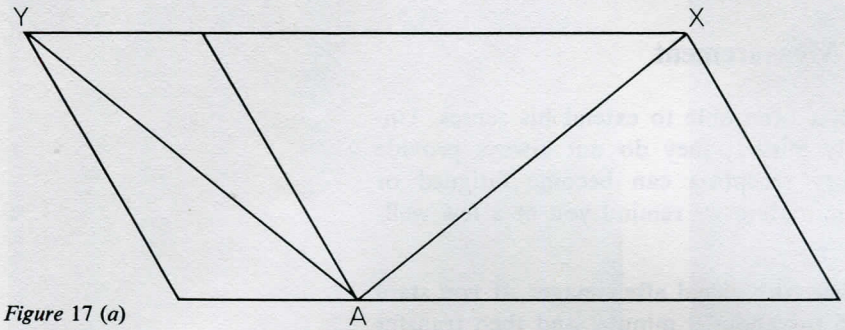


Figure 17 (*a*)

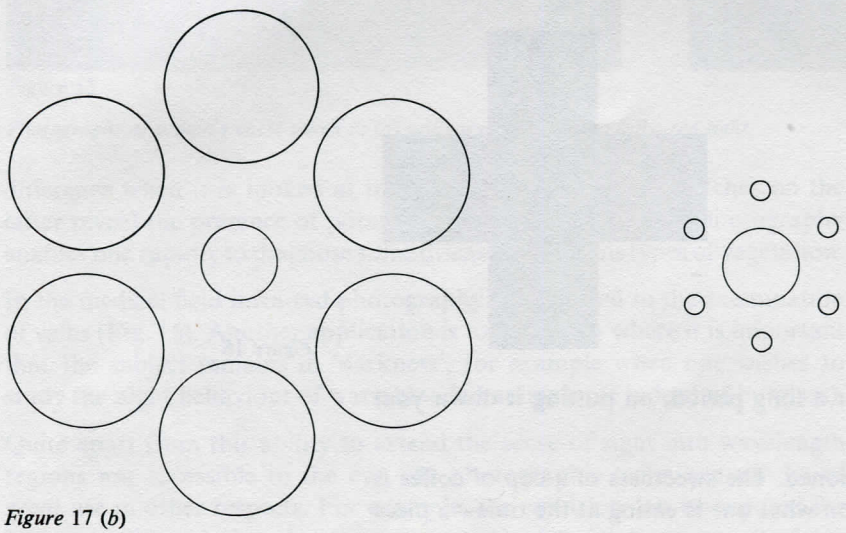


Figure 17 (*b*)

Ebbinghams' figure.

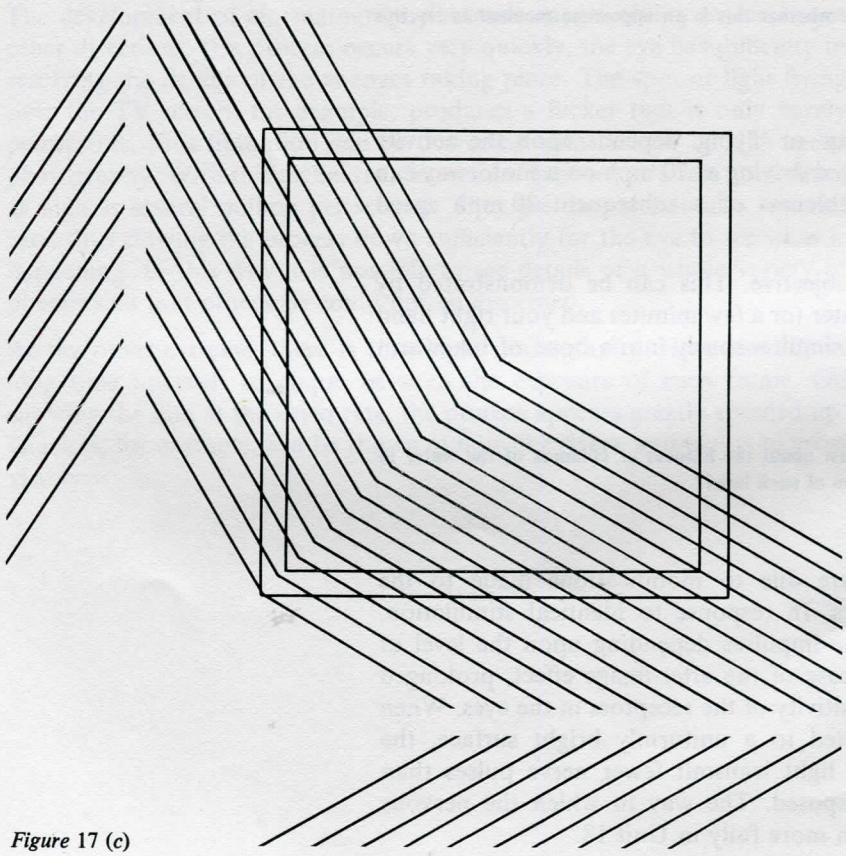


Figure 17 (*c*)

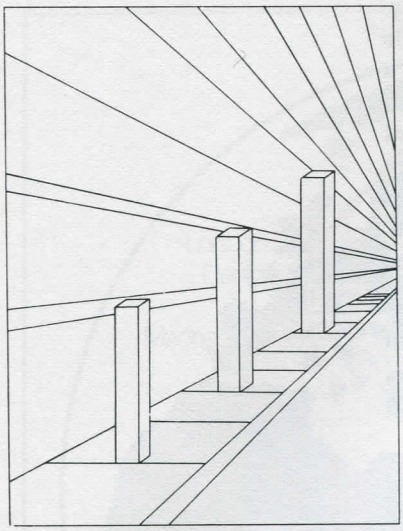


Figure 17 (*d*)

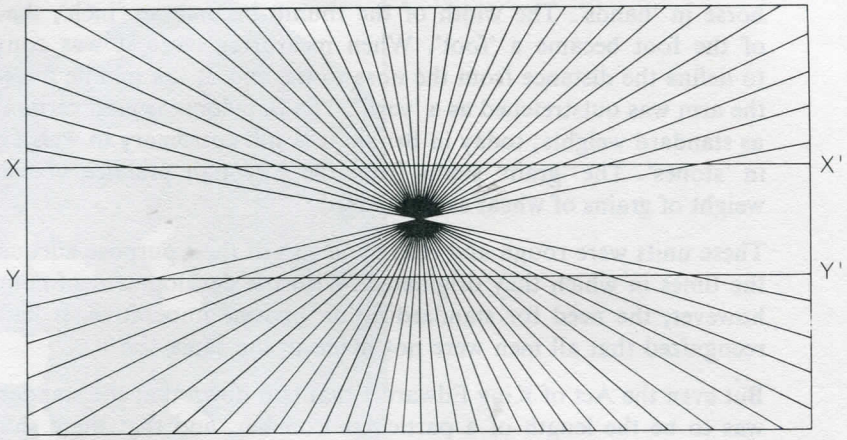


Figure 17 (*e*)

Hering's illusion.

17 (a) Are lines XA and YA of the same length?
17 (b) Are the inner circles identical?
17 (c) Is the figure a square?
17 (d) Are the upright columns identical?
17 (e) Are the lines XX′ and YY′ straight and parallel?

It is fairly well established that these latter types of illusion originate, at least primarily, in the way in which the brain organizes the information sent to it. Experiments have been performed in which the subject was able to see only part of a drawing with one eye, the rest being seen by the other eye. Under these circumstances the illusion remains. As neither eye sees the whole drawing the illusion cannot be attributed to the functioning of the eye but instead to the way in which the brain fuses together the information.

Clearly observations have to be interpreted with caution. Indeed variations in observation become even more marked when one person's experience is compared with that of another. Individuals see colours differently—this is known because some people are colour blind. Individuals hear sounds differently—some are tone deaf.

If one is to add to a common fund of knowledge information that is to be useful to others, it must be about objective reproducible experience; it must be free from the whims and peculiarities of the particular observer. Observation must develop therefore into something more quantitative —it must become *measurement*.

2.4.2 Systems of units

In measurement, certain standard *units* are agreed upon and all other quantities are expressed as multiples of these units. For example the length of a particular object, say a metal rod, may be chosen as the unit of length. The lengths of all other objects may then be related to the length of this rod. An object may be twice as long as the rod (i.e. of length two units), three-and-a-half times as long (3.5 units), a quarter as long (0.25 units), etc. Once the unit is established, measurement becomes a matter of counting.

The question arises as to what basis one should choose for the system of units. The choice is, of course, an entirely arbitrary one. Not surprisingly, therefore, many systems have appeared at different times.

Early units of length tended to be related to the size of the human body. A 'cubit' was the distance between the elbow and the end of the middle finger. The width of a man's hand was used to express the height of a horse in 'hands'. The width of the thumb became an 'inch', the length of the foot became a 'foot'. When measuring cloth it was convenient to define the distance from the nose to the end of the middle finger when the arm was outstretched as a 'yard'. The Babylonians used certain stones as standard weights; today in Britain it is still customary to weigh oneself in 'stones'. The 'grain' stems from the Egyptian practice of using the weight of grains of wheat as a standard.

These units were rough and ready and served their purpose adequately in the times in which they originated. With the development of commerce, however, the need for standardization became imperative. It had to be recognized that all men were not made to the same size!

But even the Act of King Edward I that laid down that the standard yard was to be the length of a particular iron bar, and that there should be exactly three feet and exactly 36 inches in this standard yard, did not entirely alleviate the problem. Other systems of units also became standardized. This multiplicity of systems has left a legacy of confusion—not to mention the considerable and quite unnecessary labour of converting measurements from one system to another.

Can you think of any examples of this kind of confusion?

The problem of having to convert temperatures from degrees Fahrenheit to degrees Celsius (or Centigrade), or a Continental speed limit from kilometres per hour to miles per hour, are just two examples of the trouble caused.

Fortunately, vigorous steps are now being taken to achieve uniformity. In 1960 it was formally approved by the General Conference of Weights and Measures that a system of units called Système International d'Unités (abbreviated to SI units) should be adopted. The Open University publications will in general conform to this decision.

In section 3 of *HED*, there is some further information about units. You will find it useful to refer to this at various stages during the Foundation Course and indeed throughout your degree course, if you continue to study science with us. *Don't* try to memorize it. You will be gradually introduced to the terms as you go along (some of them you will not need at all during the Foundation Year—they are there merely for the sake of completeness).

For the time being we shall concentrate on just two of these units—the *metre* as the unit of length and the *second* as the unit of time.

The unit of length: The metre was originally defined as one ten-millionth of the distance from the Pole to the Equator along a great circle† passing through Dunkirk in France and a point close to Barcelona (Fig. 18).*

This was chosen because it was a distance that could be remeasured at will (and could not be lost!). However, the measurement of it is not at all convenient. So later, two marks one metre apart were inscribed on a metal bar supported in a specified manner, at standard atmospheric pressure†, and at the temperature of melting ice.** This distance was then defined as the new standard for the metre. The bar was kept in a vault at Sèvres near Paris and copies of it were made for other countries.

* *As you can read in* The Roots of Present-Day Science (RPS), *this definition had its origins in the burst of scientific activity that stemmed from the French Revolution.*

** *The temperature must be specified because materials change their dimensions with change of temperature.*

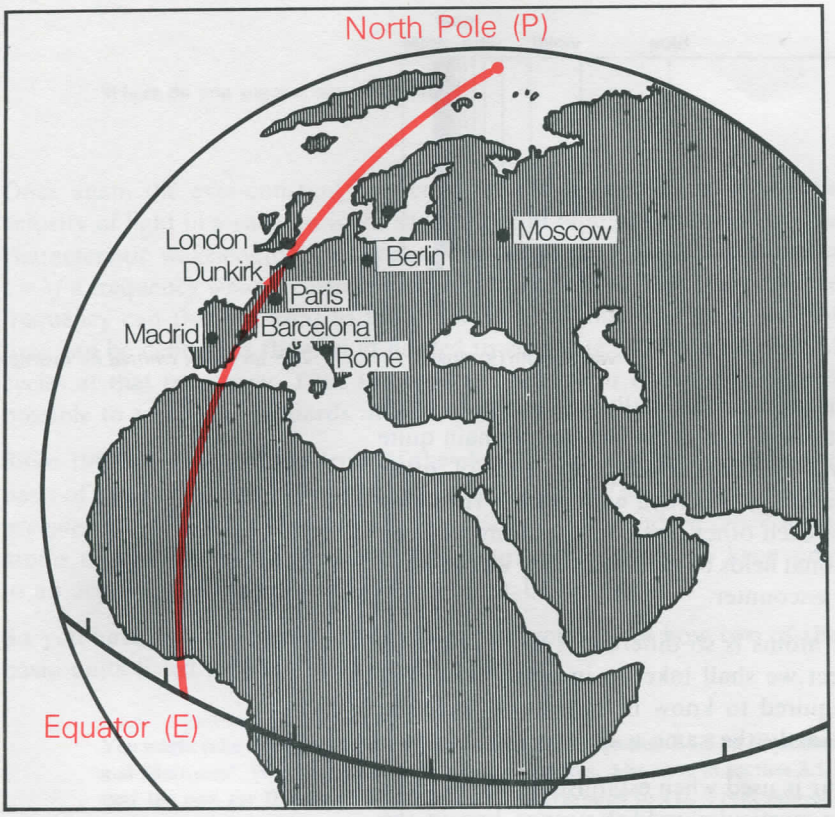


Figure 18

The metre was originally defined as one ten-millionth of the distance PE.

This was still not considered entirely satisfactory—one question that arose for example concerned the exactness of the copies. It was recognized that a still firmer basis for the definition of the metre would be desirable.

The search was for an ideal standard—one based on a property of nature that could be relied upon never to change and was moreover readily accessible to everyone without the need to make copies. It is at this point we introduce you to a remarkable feature of nature—the identical behaviour of atoms.

We said earlier that there were different types of atom; each is characterized by the number of electrons it contains. Substances built up from atoms of one type are called *elements*. Silver and bromine are two elements you came across in Appendix 1 on photography. Atoms of a given element are identical*; they have identical nuclei and identical electrons. There are a variety of possible configurations the electrons can adopt within the atom, but these possible configurations are the same for all atoms of the same element.

An electron can move from one allowed configuration to another, i.e. alter its orientation with respect to the nucleus. When making certain transitions it emits a pulse of light. The wavelength of this light is a characteristic of the transition from a particular initial configuration to a particular final configuration. These pulses of light are of extreme importance because they carry much information about what goes on inside the atom. As you will see in some later Units, by measuring the wavelength of the light emitted, one can infer which particular transition the electron has executed. Figure 19 shows a spectrum of the light emitted by atoms of the element hydrogen. In contrast to the continuous white light spectrum you will be shown in Unit 6, here you see that only certain wavelengths are present. It is called a *line spectrum*. This set of lines is characteristic of hydrogen; no matter where or when it is measured, no matter what the past history of the particular sample of hydrogen may be, this is the spectrum obtained.

** Not strictly true—there are varieties called isotopes, but these are irrelevant to our discussion. They will be dealt with in Unit 6.*

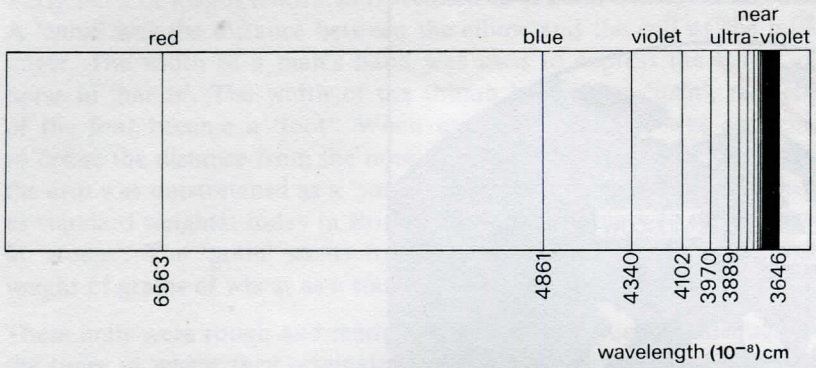


Figure 19

The spectrum emitted by hydrogen.

Why should this be so? How can atoms continually jostle and hit each other and yet, apart from the occasional loss of an electron, remain quite unchanged by their past history? This is certainly not true of other systems one can think of. If two solar systems each with planets revolving around a central sun came near to each other, the configurations of both would be distorted by the gravitational fields of the other; they would carry away with them the effects of the encounter.

The reason why the behaviour of atoms is so different is to be found in Quantum Theory. This is a subject we shall take up in detail later, but for the time being all you are required to know is that atoms have the property of always behaving in exactly the same way.

This exactly reproducible behaviour is used when establishing the modern standard of the metre. In 1960, a particular reddish-orange line in the spectrum of the gas krypton was chosen, and *the standard metre is now defined as having a length equal to 1 650 763·73 wavelengths of this light in a vacuum.** This may seem to you an extraordinarily odd definition; but in practice it is very convenient. Krypton is readily available all over the world because it is present to a small extent in the atmosphere. So anyone at anytime has now direct access to the standard—he has no need to refer to a particular object kept in a vault somewhere.

The unit of time: We turn now to the definition of a unit of time. Here, unlike length, it was not possible to take a certain quantity, define it as the standard and keep it locked away in safety. It was necessary to think of something else.

The notion of the passage of time is intimately bound up with the repeating cycle of day following night. So it first appeared natural that the standard of time should be related to the length of a day. It could be given as a fraction of the solar day (i.e. a fraction of the period of time from the moment when the sun is at its highest elevation in the sky one day, to the corresponding moment the next day). However, the length of the solar day varies throughout the year. So the unit must be based on the *mean solar day*, which is an average taken over the whole year. The division of this mean solar day into 24 hours, each hour into 60 minutes, and each minute into 60 seconds, gives the basic unit of time—*the second.*

Unfortunately the rotation of the Earth is not exactly regular—for one thing the action of the tides gradually slows down the rate at which the Earth spins on its axis, and so lengthens the mean solar day. A more consistent type of periodic behaviour is therefore required.

* *We must specify 'in a vacuum' because although a light wave does not* need *a medium in which to travel, if it* does *pass through a medium (such as the air) then it slows down. The frequency remains unaltered so we see from the relation* $v = \lambda f$ *that the wavelength must be affected if the velocity is lower. Note you are* not *required to remember the number of wavelengths.*

Where do you suggest one could look?

Once again the ever-constant properties of the atom can be used. The velocity of light in a vacuum is the same for light of all wavelengths, so the characteristic wavelength of a spectral line specifies through the equation $v = \lambda f$ a frequency which is also characteristic of the line. This reproducible frequency can then be used to fix a standard of time. The basic unit of time can be defined as the period needed to complete a certain number of cycles at that frequency. Thus the constant behaviour of atoms makes it possible to establish standards of both length *and* time.

Since 1967 a line in the spectrum of the element caesium has provided the basis of the unit of time. *The standard second is defined as that time interval occupied by 9 192 631 770 cycles of this radiation.** A standard clock using atoms of caesium as its effective 'pendulum' can be made to keep time to an accuracy of better than one second in 10 000 years.**

So you have now reached a point where you understand how two of the basic units listed in *HED*, section 3, originated.

You might take this opportunity of memorizing *HED*, section 3.1.5 on 'Fractions and Multiples'. We can then start to use the prefixes. Also note in section 3.1.3 that the unit for frequency is the *hertz*,* abbreviated to Hz. From now onwards we shall drop the use of 'cycles per second', which means the same thing.**

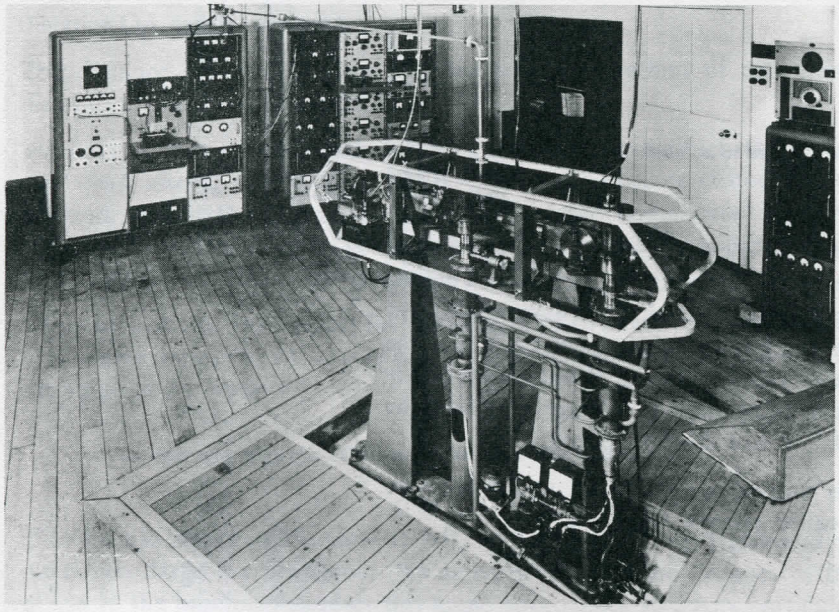


Figure 20

The world's first Caesium clock developed at the National Physical Laboratory, Teddington.

* *You are not required to remember this number.*

** *Before ordering one for the home, however, you are advised to look at Figure* 20*—it's not exactly what one normally expects a clock to look like.*

*** *Named after the German experimental physicist, Heinrich Hertz, who did much work on the nature of electromagnetic radiation.*

2.5.1 The Need to Extend the Scale of Measurement

You will have noticed that the early units of length were closely related to the size of a man's body; the unit of time to the length of a day. Such choices are not surprising. The units were devised when man was most preoccupied by his immediate observations.

But you have already seen that there is no reason why man should restrict his observations to what is directly accessible to his senses. He could also with a little ingenuity extend the scale of his measurements. But why should he want to do this? The reason is simple enough. Many of the problems that excite his curiosity or whose solution would be technologically most advantageous can only be answered by measurements made over extreme scales of length and time.

What is man made of? The answer is partly to be found in the study of tiny atomic particles. This is a world where movement is rapid and changes occur quickly—it involves, therefore, not only very small distances but also short time intervals.

How and from what did he evolve? To answer this he must find means of measuring time that stretches far beyond the span of his own existence on Earth.

These and other questions you will be tackling in this Foundation Course.

2.5.2 Measurement of large distances

How do you normally judge the distance to an object?

When judging the distance to an object, there are several clues to choose from. If your eye registers a minutely small impression of a bus, you do not immediately think 'there is a minutely small bus'. You have learnt how big a bus is normally, and so it is more reasonable to conclude that the apparent smallness is due to its being a long way off. If one object partly obscures another, it must be the nearer of the two. If there is a haze, the colours of the more distant objects are less sharply defined.

With *two* eyes, depth perception is enormously increased, and it becomes possible 'to see in depth'. This is called *stereoscopic vision*.

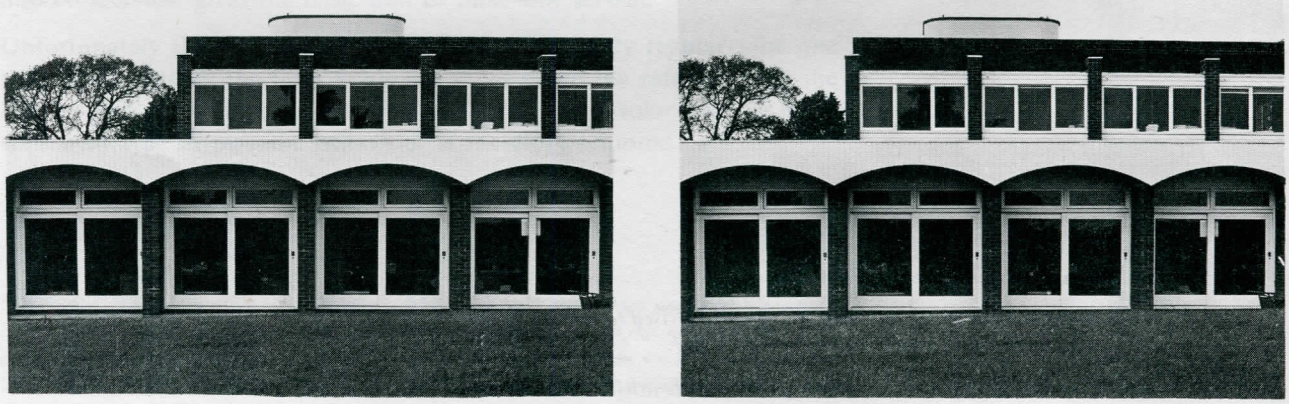


Figure 21 *A stereoscopic pair of photographs.*

The principle of stereoscopic vision is easy enough to understand. The two eyes are spatially separated, so they look at an object from slightly different vantage points. Figure 21 shows views of the library building of the Open University at Milton Keynes as seen from two such points. The relative positions of the upper and lower storey appear to have shifted, the amount of shift depending upon their distance from the camera. It is still a problem, of course, as to how the brain manages to put two pictures together to give an immediately recognizable impression of varying depths. But at least the information on relative depths is known to be there on the pictures. If presented with two pictures taken by cameras separated by a known distance, it would be possible to calculate the distances to the objects.

In Figure 22 the positions of two cameras are shown as points B and C. Directly ahead of them is an object at O.

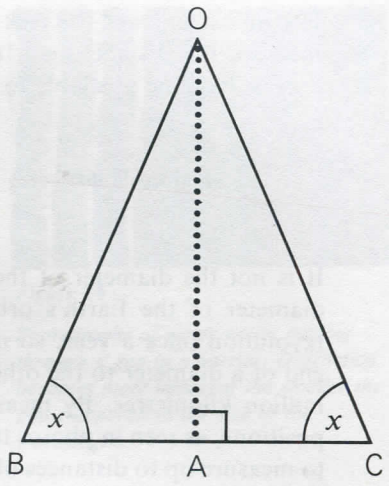


Figure 22

From simple trigonometry*

$$\frac{OA}{AC} = \tan x$$

$$\text{i.e. } OA = \frac{BC}{2} \tan x \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Suppose the cameras are situated 12 metres apart and the angle, x, is 84° for an object placed at point O in Figure 22. What is the distance to the object from the centre of the baseline?

BC = 12 metres $\therefore \frac{BC}{2} = 6$ metres

tan x = tan 84° = 9.514 (from a table giving the values of tangents of angles)

$\therefore$ From equation (2) the required distance, OA, is given by

OA = 6 tan 84° = 6 × 9.514 = 57.084

OA = 57 metres

So for each value of the distance OA there is a characteristic value of the angle x, or conversely if one knows the angle x and the distance BC, the distance OA can be deduced.

The distance BC is called the *baseline* and this method of estimating distance is known as *triangulation*.

We have chosen a particularly simple configuration in which the object was dead ahead, i.e. $\angle OBC = \angle OCB = \angle x$. When this condition is not satisfied, the mathematics becomes a little more complicated but the principle remains the same—given the angles and the baseline, the distance to the object is uniquely determined.

If you look at a table of values for the tangent of an angle you will see that, if the angle is close to 90°, a tiny change in angle makes a big difference in the value of the tangent.

In what way does this affect our estimation of distance?

If angle x in equation 2 is close to 90°, a slight error in its determination introduces large inaccuracies in the value of OA. As the distance of the object to the baseline increases this angle approaches 90°. Any particular camera set-up, therefore, is suitable for estimating distances over a certain restricted range and not beyond. The same applies to your two eyes, they are also effective only over a restricted range.

It would be a different matter if the baseline could be extended. You can see from equation 2 that, for a given angle x, the distance OA is proportional to BC. If one could for example double the baseline, one could double the range of the distance measurements.

Extending the baseline is in fact what a surveyor does. He views a distant object with his telescope, first from one position and then from another, the separation of the two positions being known. He can, of course, use very large baselines.

* *See* MAFS, *section* 4.A.1, *if necessary.*

For an astronomer the task is that much greater. Suppose he wanted to measure the distance to a star he would have to look around for the longest baseline available to him.

What would that be?

It is not the diameter of the Earth, as you might think at first, but the diameter of the Earth's orbit around the Sun. The Earth completes a revolution once a year, so in a period of six months it moves from one end of a diameter to the other (Fig. 23). The length of this baseline is 300 million kilometres. By measuring how much stars appear to shift their positions, as seen in photos taken from these two vantage points, he is able to measure up to distances of about 5×10^{15} km. Triangulation is therefore an effective means of distance estimation for the astronomer—it essentially increases the distance between his eyes.

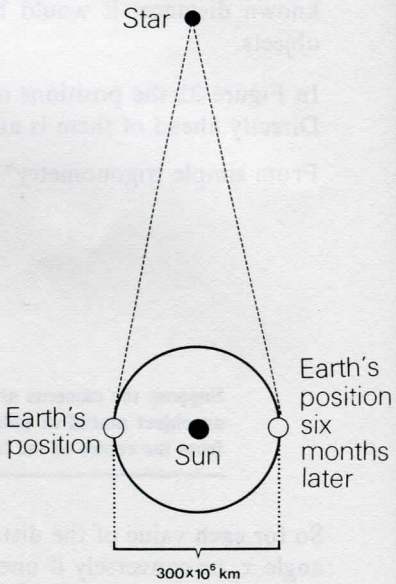


Figure 23

Triangulation using the diameter of the Earth's orbit as base line.

2.5.3 Measurement of small distances

It is obviously not possible to go on subdividing a length indefinitely without reaching a stage where it is no longer possible to see the subdivision. The object could of course be brought closer to the eyes so that it looked larger, but as you know there is a limit to how near an object can be brought for it still to look clear and sharp. So instruments must be used to aid the eye. The simplest is the *magnifying glass* and a more complicated one—*the microscope*.

At this point we check your knowledge of simple optics. Do you need to study either or both of the following Appendices?

Appendix 4 (Red) on Lenses:

1. **Do you understand the following terms:**
   - **(i) a ray?**
   - **(ii) refraction?**
   - **(iii) an image?**
   - **(iv) a lens?**
   - **(v) lens aberration?**
2. **When a ray of light passes through a slab of glass, the angle through which it is deviated depends on four factors. Can you name all four?**

If you fail to answer any item study Appendix 4.

Appendix 5 (Red) on The Optics of the Eye.

1. **Can you explain the formation of an inverted image on the retina?**
2. **Can you explain how our eyes are able to focus on objects at different distances?**

If you fail on either question study Appendix 5.

NOW (WHETHER OR NOT YOU HAVE WORKED THROUGH APPENDICES 4 and 5) TURN TO APPENDIX 2 ENTITLED 'THE PRINCIPLE OF THE MICROSCOPE' AND READ SECTION 2. THIS IS WHITE-PAGE MATERIAL.

A great deal of expertise goes into the design and manufacture of microscopes in order to produce highly magnified, sharply focused images. But eventually optical instrument makers come up against a barrier—one imposed by the nature of light itself.

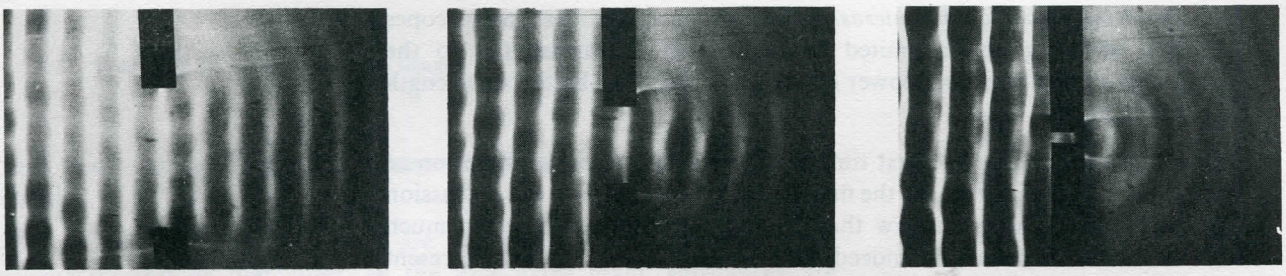


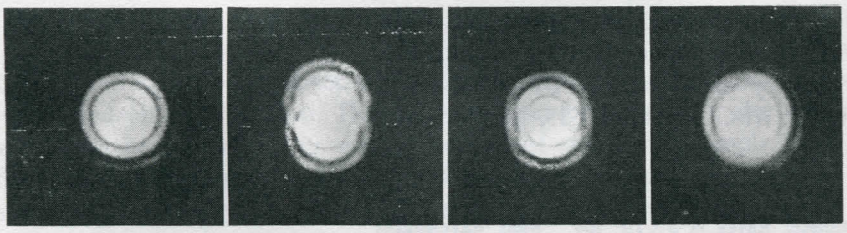
Figure 24

Photographs of water waves passing through a gap in a barrier. Diffraction becomes more apparent the smaller the gap compared to the size of the wavelength.

When discussing the optics of the microscope (in Appendix 2), we drew straight lines in the diagrams to denote light rays. Light however is not so well-behaved. You saw earlier that it had a wave nature. Waves, when they pass close to obstacles, can change direction. To illustrate this, we show in Figure 24 water waves of various wavelengths passing through a gap in a barrier. The waves are travelling from left to right, so this is the initial direction of the rays (you are to imagine them at right-angles to the wavefronts). When the wavelength is small compared with the size of the hole, the waves passing through it continue mainly in this direction. But when the wavelength of the wave is about equal to, or greater than, the size of the hole, you can see that the wave bends round the corner to a considerable extent—it spreads into regions on either side that you would perhaps expect to be in 'shadow': this characteristic behaviour of waves is called *diffraction*.

We said earlier (p. 14) that an *object* comparable in size to that of a wavelength, scatters only a *little* of the incident wave—most of it continues on undisturbed. Now we are saying that a *hole* comparable in size to a wavelength produces a *large* disturbance with much of the wave being deviated.
How do you reconcile these two observations?

There is no contradiction. One would expect an object to cast a shadow. But if the object is small (as in Fig. 5) the waves diffracted round its edges spread into the region behind the object to such an extent that the shadow is almost non-existent. If the intensity of the on-going waves in this region is so little reduced, then clearly not much of the wave is left to be scattered into other directions.



(*a*) (*b*) (*c*) (*d*)

Figure 25

Diffraction patterns produced by a lens: (a) image of a single point object; (b), (c) and (d) images of two point objects being brought closer together.

When light waves pass through a lens set in a microscope, they too are passing through an aperture. The rim of the lens is acting as the boundary of a hole—any light not striking within the rim of this hole is prevented from passing through to the other side by the opaque material of which the microscope is constructed. The light will, therefore, be slightly diffracted (only slightly because the diameter of the hole is very much larger than a wavelength). The expanse of glass straddling the hole serves to concentrate the rays; but diffraction spreads the light out a little and so gives rise to a blurred image rather than a point image. Figure 25 (a) shows the sort of effect produced. There is a central bright region surrounded by alternating rings of light and dark (why the image takes on this particular form will be explained later in Unit 28). Figure 25 (b) shows images of two well separated objects showing similar patterns. In Figures 25 (c) and (d) they are brought closer together. In the latter case it is more or less impossible to decide whether there really is a second object. Thus, if the distance between two objects is smaller than some minimum value, it is no longer possible to resolve them. *This minimum separation is known as*

the resolving power of the microscope. In practice, the best microscopes have a resolving power limited to about half a wavelength. So the limitation on the resolving power of a microscope is set by the wavelength of the light used.

This, of course, is not the first time you have come across a limitation of this type. You will remember the first was in connection with our discussion of ultrasonics. There you saw that objects could not be detected much smaller than a wavelength. Indeed, another way of looking at the present problem is to say that, once the size of the specimen becomes smaller than the wavelength of the light that is illuminating it, it becomes very inefficient at reflecting light into the microscope.

One attempt to remedy this, of course, would be to increase the size of the objective lens. This would increase the angle, ϕ (phi), subtended† by the lens at the object (Fig. 26), and would allow the lens to collect a greater fraction of the light scattered by the object. But the improvement to be gained this way is limited.

What else could you suggest might lead to an improvement in resolving power?

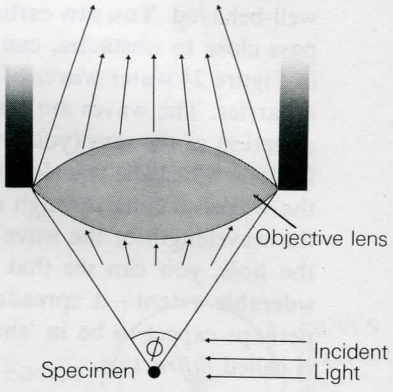


Figure 26

Diagram illustrating the angle subtended by the objective lens at the specimen being viewed under a microscope.

Clearly an alternative approach is that which was adopted in the case of sound waves, i.e. to reduce the wavelength of the illumination. Does electromagnetic radiation reflect from objects more efficiently if the wavelength is smaller? The answer is—yes. The blue colour of the sky and the redness of sunsets are examples of everyday phenomena produced by this behaviour of light.

At this stage you should do the home experiment (Unit 2, Experiment 1) which goes into this topic further.

A further example of the dependence of the scattering of electromagnetic radiation on wavelength is to be found in the reception of TV waves. In this country we transmit on both VHF (**V**ery **H**igh **F**requency) and UHF (**U**ltra **H**igh **F**requency). On VHF, for example, BBC 1 (Channel 1, London) is transmitted at a frequency of 45 MHz, ITV (Channel 9, London) on 195 MHz; and on UHF, BBC 2 (Channel 33, London) at a frequency of 573 MHz.

If the velocity of electromagnetic waves is 3×10^8 ms^{-1} what wavelengths would these three waves have?

For the first broadcast channel,
$f=45$ MHz$=45\times10^6$ s^{-1}.
From the equation $v=\lambda f$:

$$3\times10^8=\lambda\times45\times10^6 \quad \therefore \lambda=\frac{3\times10^8}{45\times10^6}$$

$$=\frac{300}{45}=6.67 \text{ m}$$

Similarly for the other channels the wavelengths can be found to be 1.54 m and 0.523 m.

The difference in these wavelengths makes it easier to scatter UHF waves compared to the others and so for these there tend to be more 'shadows', or pockets of bad reception.

If you have a conveniently placed indoor TV aerial you might like to get a feel for how easy or difficult it is to shield the aerial from the waves. Try holding or suspending objects of various sizes (preferably metal ones like saucepans, aluminium foil, a dustbin lid, etc.) close to the aerial in a variety of positions. Watch their effect on the picture. Try it for UHF and VHF. (Note that if you are not careful the shielding effect produced by your hand, arm, etc., may swamp the effect you are looking for.)

Thus, in order to see still smaller objects under a microscope it would appear that the answer is to go to smaller wavelengths. Consequently some microscopes use electromagnetic radiation with wavelengths just beyond the visible region, namely ultra-violet light.

What could you use to view the image?

Such microscopes, fitted with cameras as detectors, have proved quite useful.

At still shorter wavelengths there are X-ray microscopes. But here one encounters serious difficulties. In the first place many substances are transparent to X-rays and so scatter little radiation. Secondly, as mentioned in Appendix 4, the refraction undergone by an electromagnetic wave is dependent upon wavelength; for X-rays the refracting effect is very small indeed and so they prove difficult to focus. Some microscopes have been developed which dispense with lenses and focus the rays instead with mirrors made of curved crystals. Even so, the practical difficulties of getting images of good intensity and contrast restrict the performance of the instrument to well below the theoretical limit based on wavelength considerations alone. In fact their ability to resolve small distances is not greatly different from that of microscopes using light in the visible region.

Thus the possibilities of microscopes using electromagnetic radiation as their source of illumination appear to be exhausted. In order to extend still further the measurement of small distances, a radically different technique is called for. We now describe the electron-microscope—a technique that 'illuminates' the object with a beam of electrons.

So far you have learnt that electrons are tiny, negatively charged particles and that they are one of the constituents of atoms. Basically the idea behind the electron-microscope is that a beam of electrons is directed at the specimen. The specimen, in scattering the electrons, casts a large 'shadow' that can be viewed. However, we have so far only spoken of the electron as being attached to an atom; we must therefore describe, before going any further, how it is possible to obtain a beam of free electrons.

What do you think holds electrons to the atom?

It is the attractive force exerted by the positive charge on the nucleus that holds electrons to the atom. The different types of atom have electrons which can be detached with varying degrees of ease. For example, it is found to be comparatively easy to detach electrons from atoms of metallic substances such as iron or silver. This being the case, when these atoms are pushed close together, as they are in a solid, the outermost electrons in the atoms come within the sphere of influence of the electrical fields of neighbouring atoms. Some electrons are then attracted away from their parent nucleus to a neighbouring one. This may seem to contradict what we said earlier, namely that a charged body brought close to an atom would not experience any net force due to the atom. Previously we presented the argument that the forces exerted by the electrons on the one hand and the nucleus on the other would cancel each other out. It all depends on what we mean by 'close'. If the approach of the outermost electron is not too close (position A in Fig. 27), what we said earlier is true—its distance from the electrons of the neighbouring atom is about

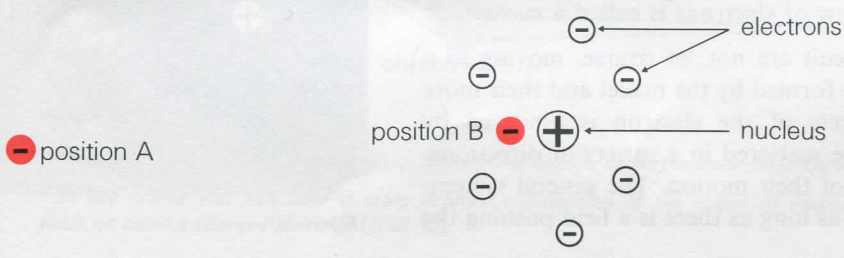


Figure 27

the same as its distance from the nucleus, so the forces are equal in magnitude but act in opposite directions. If the approach is exceedingly close however (say position B in Fig. 27), the difference in these distances becomes as large as the distances themselves and so can no longer be ignored—the forces exerted by the electrons and by the nucleus no longer cancel and a net force results. This is what happens in a solid. A loosely bound electron from one atom is able to get right in among the electrons of a second atom and thereby be attracted away from its original nucleus towards the neighbouring one.

It can then be attracted to the next one and so on. So a metal essentially consists of a lattice of nuclei surrounded by some electrons that are bound closely and tightly to their parent nuclei, and others that are free to wander around throughout the specimen. These nomadic electrons are called *conduction electrons* and are responsible for the way metals are good conductors of heat and of electricity.

Conduction electrons normally stay within the confines of the specimen of metal. This is because, when any tend to drift away, the specimen as a whole is left with an unbalanced positive charge that promptly attracts them back again.

The motion of the conduction electrons is usually quite random—there are as many moving in one direction as another, and throughout the specimen they are evenly distributed. If, however, the metal is placed in an electric field, the effect of the field will be felt throughout the interior of the metal (Fig. 28). The electrons within the metal will experience a force and those that are free to move in response to this force will do so. The result is that, in addition to the random motion, there will also be a general systematic drift in one direction due to the field (Fig. 29). This systematic movement of electrons cannot continue indefinitely because they come up against the surface of the specimen and this acts as a barrier to further movement (Fig. 30). If, however, the metal is in the form of a closed loop, say a length of wire, the movement of electrons can be maintained for as long as there is an electric field acting around the loop.

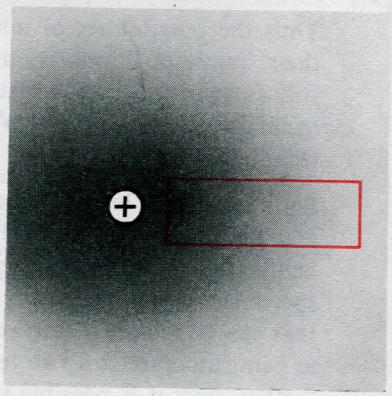


Figure 28

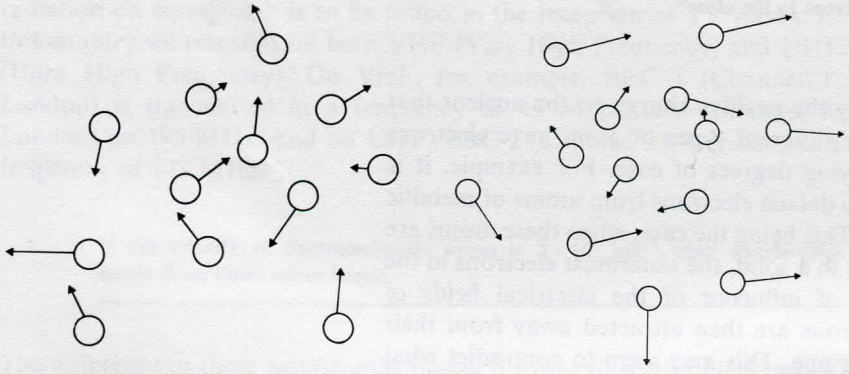


Figure 29 (*a*)

No electric field: random motion.

Figure 29 (*b*)

Electric field: random motion plus systematic drift.

A *battery* is a device for maintaining such an electric field. By connecting the ends of the wire to the terminals of the battery a *circuit* is established (Fig. 31). Electrons derived from the chemical reaction in the battery can flow through the wire, into the battery at one terminal, out at the other, and round the wire again. Such a flow of electrons is called a *current*.

The electrons passing round the circuit are not, of course, moving in a vacuum; they move through a lattice formed by the nuclei and their more tightly bound electrons. The progress of the electron is hindered by impurities in this lattice, and they are scattered in a variety of directions. This adds to the random component of their motion. The general systematic drift can only be maintained for as long as there is a field pushing the electrons along.

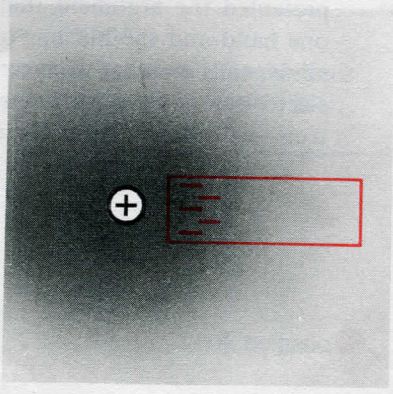


Figure 30

So the application of a field not only produces a current but also, because of this scattering, an increase in the random motion of the conduction electrons. And the faster on the average this random motion, the hotter the specimen is said to be. This is precisely what happens when one switches on an electric fire. A current is made to flow through a wire (called an element), this heats up and passes the heat out to the room.

When a metal becomes very hot indeed (either by being heated electrically as above or by some other means) an interesting thing happens. The conduction electrons bounce around inside so fast that some of them are ejected completely from the specimen. To be sure, the unbalanced charge on the metal tries to pull them back, but if the speed of ejection from the surface is sufficiently great the electrons can get entirely away. It is rather like a projectile being launched from the surface of the Earth; normally gravity pulls it back, but if the speed of a rocket is sufficiently great it can escape the pull of gravity.

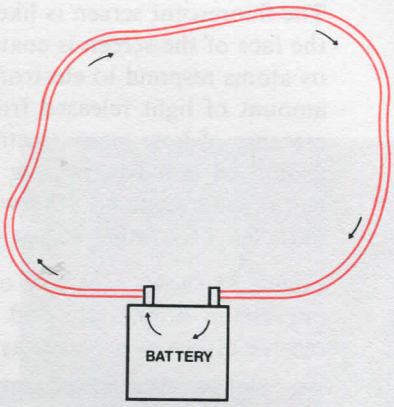


Figure 31

This effect provides a simple means for obtaining free electrons. Such electrons can be enormously useful. One of their applications is to be found in the *electron microscope*. Figure 32 (a) shows a schematic layout of such a microscope. F is a wire filament that is heated until it glows; it continuously gives off a cloud of electrons. These are immediately subjected to a powerful attractive electric field exerted by the charged metal plate M. The plate has a hole in it through which the accelerated electrons can pass. They are then directed on to the object that is to be viewed. The electrons scattered from the object are then focused on to a fluorescent screen S. The focusing is produced by magnetic fields which serve to alter the paths of the electrons in much the same way as optical lenses refract light in an ordinary optical microscope.* Comparison between Figures 32 (a) and 32 (b) shows how similar are the optics of the two types of microscope.

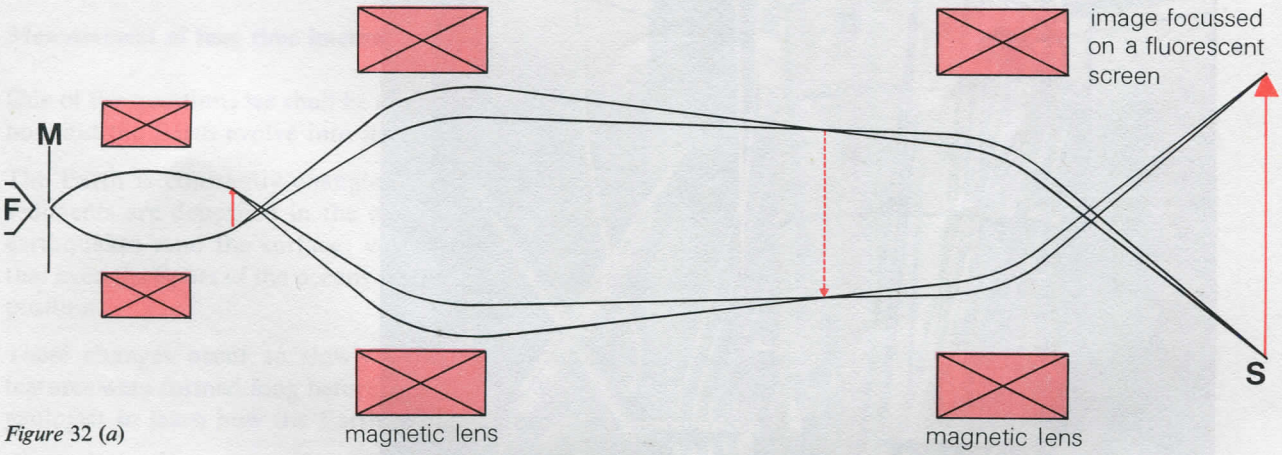


Figure 32 (*a*)

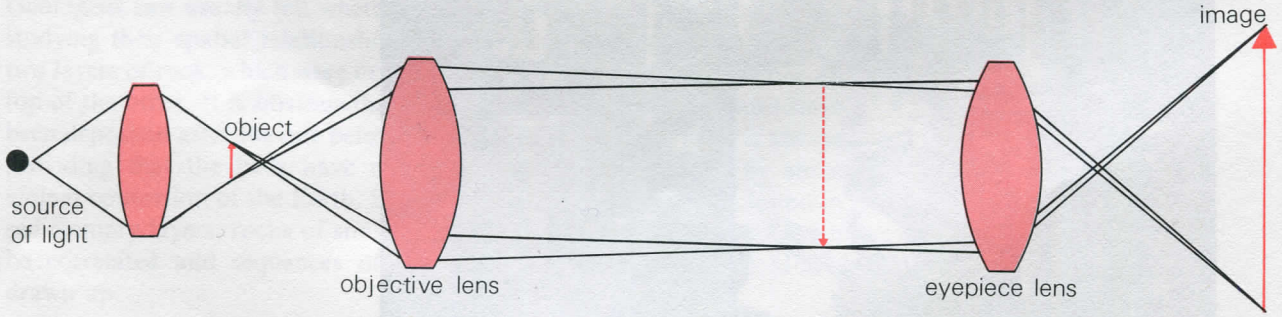


Figure 32 (*b*)

A comparison between the optics of (*a*) *an electron microscope and* (*b*) *an optical microscope.*

** In due course you will have a more detailed explanation of the action of magnetic fields on moving charged particles* (*Unit* 32).

The fluorescent screen is like the one in your television set. The inside of the face of the screen is coated with a substance having the property that its atoms respond to electron bombardment by emitting visible light. The amount of light released from the various parts of the screen is then a measure of how many electrons are striking that region of the screen. It should be noted in passing that the interior of the electron microscope has to be evacuated otherwise the electrons would scatter from the air particles. (The same is true of a television tube.)

Figure 33 shows a picture of a modern electron microscope; it gives an impression of the size and complexity involved. Such instruments can resolve distances as small as about 2 or 3×10^{-10} m.

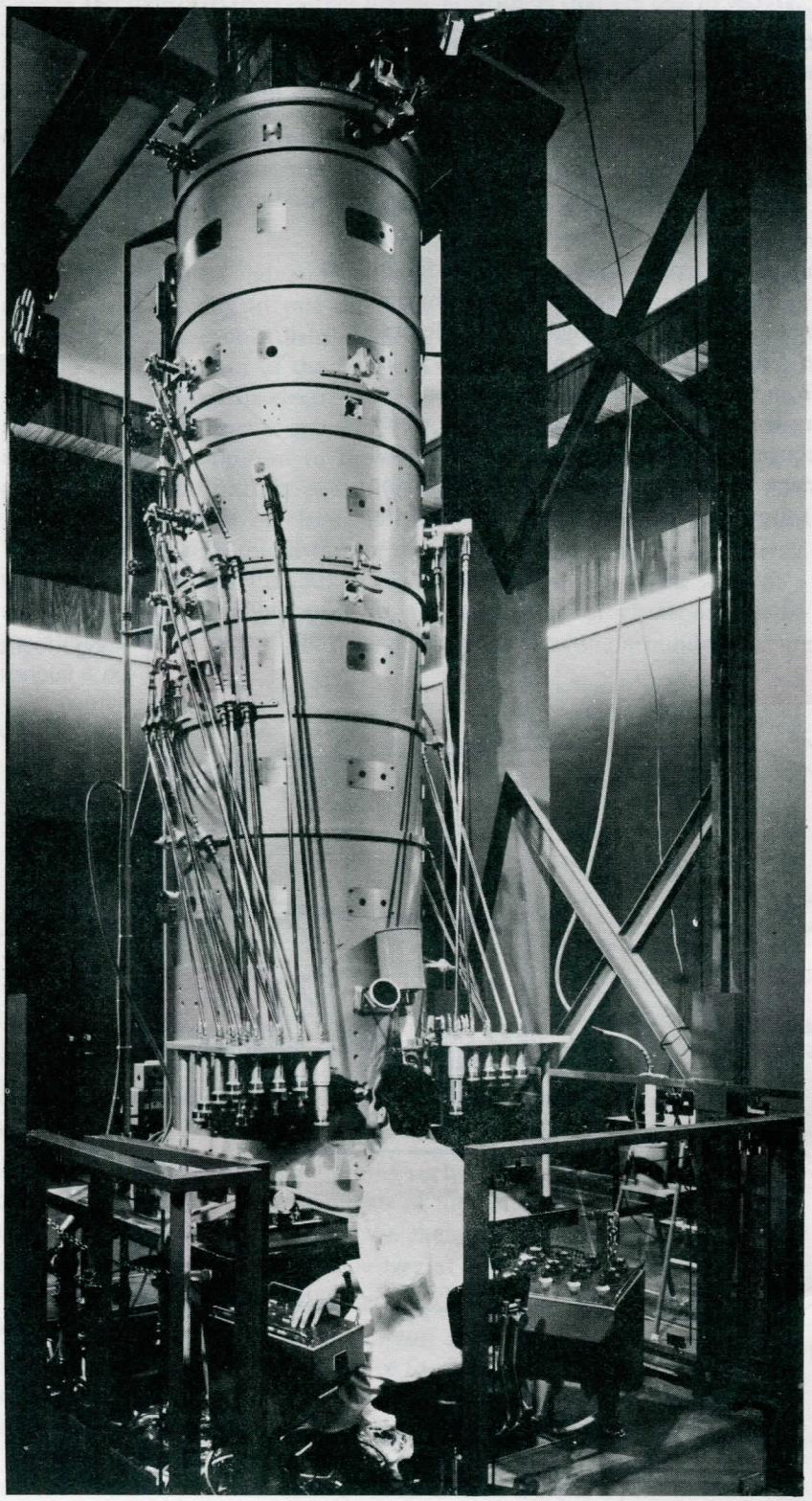


Figure 33

A large electron microscope at the Laboratoire d'Optique Electronique, Toulouse, France.

This is a very remarkable feat, but why, you may ask, stop there? If one is illuminating the specimen with a beam of particles, the type of limitation inherent in the use of waves has surely been overcome.

Now look at Figure 34. It is a photograph taken with an electron beam. What do you notice about it?

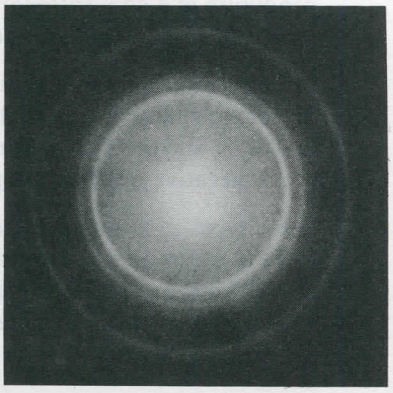


Figure 34

A photograph taken with a beam of electrons. You will see how this was done in the television programme of Unit 29.

In Figure 34 you come across a type of pattern you have seen before. With the ordinary optical microscope you saw that under high magnification the images became blurred and surrounded by circular bands.

What was the name given to the phenomenon that caused this?

So you see that electrons, which we have so far called 'particles', exhibit diffraction just like waves. This sets a limit to the ultimate resolving power that can be achieved with an electron microscope.* This ultimate limit is governed by the wavelength and this in turn is found to depend upon the momentum** of the electrons; the higher the momentum—the smaller the wavelength. The relationship is in fact a very simple one:

$$\lambda = h/p$$

Where λ is the wavelength of the electron, p is its momentum and h is a constant factor. So theoretically at least there is no limit to the resolving power achievable—ever smaller distances may be measured by using ever more energetic electrons. In practice, however, as in most modern sciences, there is the problem of finance—machines capable of accelerating electrons or other particles to great energies are expensive.

Later in Unit 29 we shall return to this topic of electron waves.

2.5.4 Measurement of long time intervals

One of the questions we shall be discussing in this Foundation Course is—how did the Earth evolve into its present state?

The Earth is constantly changing: rocks are eroded by rain and wind; sediments are deposited in the oceans, lakes and rivers; coastlines alter; earthquakes rend the surface; volcanoes build mountains; and it seems that even the floors of the oceans slowly spread and continents change their position.

These changes occur so slowly that virtually all of the Earth's major features were formed long before man existed. It is one of the tasks of the geologist to learn how the Earth has developed since its creation.

Geologists can usually tell whether one rock is younger than another by studying their spatial relationship. The simplest example of this is when two layers of rock, which were originally two layers of sediment, lie one on top of the other. It is obvious that the upper layer of sediment must have been deposited after the one below and is therefore younger; this is always providing that the beds have not been turned upside down by some violent contortion of the Earth. Similarly, by studying the fossils found in sedimentary layers, rocks of similar age in various parts of the world can be correlated and sequences of rocks, known as stratigraphic tables, drawn up.

* *We say 'ultimate' resolving power because in practice the actual resolving power of the instrument is more likely to depend upon its focusing properties.*

** *The momentum of a particle is defined as the product of its mass and velocity.*

However, neither of these types of investigation give any *absolute* idea concerning the age of rocks. To do this the geologist must find a naturally occurring process that takes place at a well-defined rate. This could then be used as the basis of a 'clock'—much in the same way as the swing of a pendulum in a conventional clock allows one to count off the seconds at a steady rate. Radioactivity provides such a process.

In order to discuss radioactivity, we must return to the study of the atom. So far you have learnt that the atom consists of a massive nucleus and that it is surrounded by tiny electrons. The electrons together carry as much negative electric charge as the nucleus carries positive charge.

Some types of atom are capable of spontaneously transforming into other types of atom. In these transformations the central nucleus emits radiations of one kind or another—often a positively or negatively charged particle. This process is called *radioactivity*, and one speaks of nuclei undergoing radioactive *decay*.

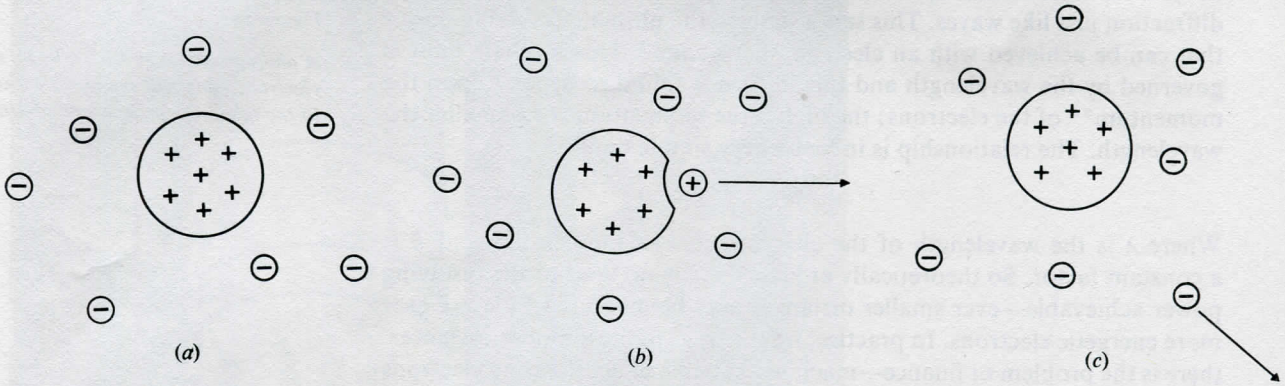


Figure 35

Radioactive decay in which a positively charged particle is emitted by the nucleus: (a) An atom with seven negatively charged electrons and a nucleus carrying seven units of positive charge; (b) the nucleus undergoes radioactive decay by emitting a positively charged particle, leaving six units of positive charge on the nucleus compared to seven units of negative charge on the electrons; (c) one of the electrons becomes detached, so leaving an atom with an overall charge of zero once again. Note that the atom in (c) differs from that in (a) both in the charge on the nucleus and in the number of electrons—it is an atom of a different element.

In all transformations in nature it is found that the total amount of electric charge remains constant—or more strictly speaking, the amount of positive charge less the amount of negative charge remains fixed. This is called the *law of conservation of electric charge*. The emission of a charged particle from the nucleus, as depicted in Figure 35, therefore, alters the charge on the nucleus and consequently gives the atom an overall electrical charge. This in turn leads to electrons being gained from or lost to the surroundings in order to rebalance the charge on the atom as a whole. The final result is that an atom of one element has transformed into an atom of another element.

You need not worry at this stage about the details—what particles are emitted from the nucleus, what elements are involved in the transformations, etc. We shall deal with these aspects in later course units. We shall concern ourselves here with the general principle underlying *radiometric dating*, and in so doing introduce you to the broad class of phenomena described by what are called *exponential functions*.

Suppose we have nuclei of a certain type, X, and these are unstable and transform to nuclei of another type, Y. These transformations occur at a certain rate. The time taken for half the X nuclei to decay to Y nuclei is called the *half-life* of the transformation. We can denote it by the symbol, $T_{\frac{1}{2}}$, and use it as a convenient measure of the speed with which the transformation proceeds.

The half-life not only specifies the time taken to reduce an initial sample of X-nuclei to a half, but applies equally to the behaviour of any subsequent sample. Supposing the number of X nuclei present initially to be N, one would expect $N/2$ to remain after a time $T_{\frac{1}{2}}$. This number will itself be reduced by a half in a further period $T_{\frac{1}{2}}$. Thus after $2T_{\frac{1}{2}}$ one expects $N/4$ X nuclei to be left; after $3T_{\frac{1}{2}}$, $N/8$ nuclei, etc.

The numbers remaining in general will not be *exactly* equal to $N/2$, $N/4$, $N/8$, etc. This is because they will be subject to statistical fluctuations. For any particular nucleus there is no way of predicting when exactly the transformation will take place. One is only able to speak in terms of probability; thus the probability, or chance, of a particular nucleus decaying in a time period equal to a half-life is 50 per cent. As it is impossible to predict for a given nucleus when it will decay, it is clearly also impossible to know how many nuclei in a given sample will have decayed in a half-life. All one can say is that it should be about a half and the bigger the sample the closer it is likely to be to a half.

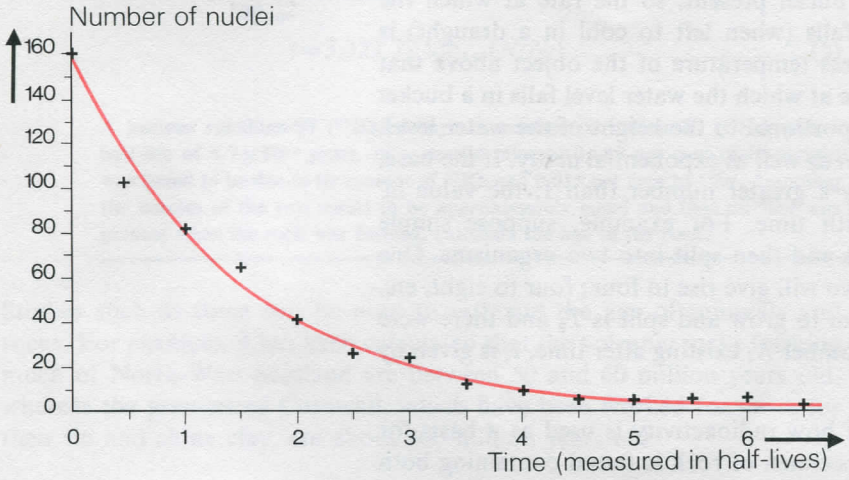


Figure 36

An exponential decay curve.

In Figure 36 we show the results of an experiment in which there were initially 160 nuclei. The smooth theoretical curve has been drawn to show that after one half-life we expect 80 nuclei to be left, after two half-lives we expect 40, etc. This theoretical curve is called an exponential decay curve. Following the same general trend, but not necessarily lying exactly on the curve, are the experimental readings.

You should now do the home experiment (Unit 2, Experiment 2).

Would you expect $T_{\frac{1}{2}}$ to depend upon the past history of a given sample of X nuclei? For example, would you expect it to depend upon how long the nuclei have already existed?

No. Since, at any given stage, it takes the same period $T_{\frac{1}{2}}$ to reduce the size of the sample by a further half, the rate at which the reaction occurs in no way depends on the past history of the particles. $T_{\frac{1}{2}}$ is also unaffected by extremes of temperature or pressure, to which radioactive materials occurring naturally in the Earth may have been subject in the past. Such extreme conditions affect only the electron configuration in an atom; much greater energies are involved in any alteration to the nucleus itself.

From the above we can deduce a formula relating the original number of X nuclei, N, to the number, N_t, remaining after any chosen time interval, t. This is:

$$N_t = N(\tfrac{1}{2})^{t/T_{\frac{1}{2}}} \quad \ldots\ldots (3)$$

Check that this formula gives the same results as before by substituting successively the time intervals mentioned earlier, viz. $t=T_{\frac{1}{2}}$, $t=2T_{\frac{1}{2}}$, $t=3T_{\frac{1}{2}}$. Refer if necessary to *MAFS*, section 5.C.6.
If there are initially 1 600 X nuclei and the half-life is 3 hours, how many X nuclei would be expected to remain after 12 hours?

$N=1\,600$, $t=12$, and $T_{\frac{1}{2}}=3$, so from equation (3):
$N_t=1\,600\,(\frac{1}{2})^{12/3}$
$=1\,600\,(\frac{1}{2})^4$
$=1\,600\,(\frac{1}{16})=100$
$\therefore$ the number remaining is 100.

The expression on the right-hand side of equation 3 is called an *exponential function.*

Can you guess why?

This is because the quantity that is changing (i.e. the variable) which in this case is the time, t, appears as an exponent.*

Though we have so far spoken only of radioactivity, you should note that there are many types of physical behaviour described by expressions of this form. Just as the number of nuclei decaying in a given time is directly proportional to the number of nuclei present, so the rate at which the temperature of a heated body falls (when left to cool in a draught) is directly proportional to the excess temperature of the object above that of its surroundings. Also, the rate at which the water level falls in a bucket with a hole in the bottom is proportional to the height of the water level. One may have exponential *growth* as well as exponential *decay*. If the base, '$\frac{1}{2}$', in equation 3 is replaced by a greater number than 1, the value of the expression will increase with time. For example, suppose simple organisms grow to a certain size and then split into two organisms. One organism will give rise to two; two will give rise to four; four to eight, etc. If the time taken for an organism to grow and split is T_2 and there were initially N organisms, then the number N_t existing after time, t, is given by $N_t=N2^{t/T_2}$.

We now turn to the question of how radioactivity is used as a basis for time measurement. Suppose a specimen of rock is found containing both X and Y atoms, and that there are reasons for assuming that all the Y atoms have been created as a result of the disintegration of X atoms (i.e. no Y atoms were present when the rock was formed), and also that the ratio of X atoms to Y atoms has not been affected by any preferential loss of either X or Y to the surroundings. (This is known as a closed system.) Knowing the characteristic half-life for the reaction X→Y (from a separate experiment carried out on another specimen of X), and having determined by analysis the numbers of the atoms of both types present in the rock, one can calculate what length of time was required to produce the ratio of X to Y.

As a simple illustration, suppose the reaction to be characterized by a half-life of 5 500 years and that for every X atom there are now 15 Y atoms. This means that originally there were 16 X atoms where now there is only one. The other 15 X atoms have decayed to Y atoms in a certain time we shall call t. Then in equation 3, $N_t=N(\frac{1}{2})^{t/T_{\frac{1}{2}}}$ .....(3)

$N_t=1$, $N=16$, $T_{\frac{1}{2}}=5\,500$ years
$\therefore \quad 1=16\,(\frac{1}{2})^{t/5500}$
hence $\frac{1}{16}=(\frac{1}{2})^{t/5500}$ But $\frac{1}{16}=(\frac{1}{2})^4$
$\therefore\ (\frac{1}{2})^4=(\frac{1}{2})^{t/5500}$
hence $4=t/5500 \quad \therefore\ t=4\times 5\,500$ years
$t=22\,000$ years

* *An exponent is the number indicating the* power *of a quantity. For example the exponent of* x *in* x^4 *is* 4. (*See* MAFS, *section* 1.A.1.)

In a more complicated example (one in which t is not arranged to be a convenient integral multiple† of $T_{\frac{1}{2}}$) you would have to solve equation (3) in a more general way:

$$N_t = N(\tfrac{1}{2})^{t/T_{\frac{1}{2}}}$$

$$\therefore \quad N_t/N = (\tfrac{1}{2})^{t/T_{\frac{1}{2}}}$$

$$\therefore \text{ Inverting } \quad \frac{N}{N_t} = 2^{t/T_{\frac{1}{2}}}$$

Taking logarithms of both sides (if necessary refer to *MAFS*, section 1.B).

$$\log (N/N_t) = \frac{t}{T_{\frac{1}{2}}} \log 2$$

$$\therefore \quad t = \frac{T_{\frac{1}{2}} \log (N/N_t)}{\log 2} \qquad \text{.....(4)}$$

If you use logarithms to the base 10

$$\log_{10} 2 = 0.3010,$$

$$\frac{1}{\log_{10} 2} = 3.322$$

$$t = 3.322\ T_{\frac{1}{2}} \log_{10} (N/N_t) \qquad \text{.....(5)}$$

A nucleus rubidium-87 (^{87}Rb) decays to another, strontium-87 (^{87}Sr) with a half-life of 4.7×10^{10} years. In a certain mineral 0.680 per cent of its weight was found to be due to its content of ^{87}Rb and 0.015 per cent to ^{87}Sr. Assuming the masses of the two nuclei to be approximately equal and that no ^{87}Sr was present when the rock was formed, calculate the age of the rock.

As the nuclei can be assumed to be of equal mass the ratio of the percentage weights is the ratio of the number of nuclei present.

$$N/N_t = \frac{0.680 + 0.015}{0.680} = 1.022$$

$$\log_{10}(N/N_t) = \log_{10} 1.022 = 0.0095$$

From equation (5), using $T_{\frac{1}{2}} = 4.7 \times 10^{10}$ years

$$t = 3.322 \times 4.7 \times 10^{10} \times 0.0095$$

$$\approx 1.5 \times 10^{9} \text{ years}$$

Studies such as these can be used to estimate the age of minerals and rocks. For example, it has been calculated that the volcanic rocks forming much of North-West Scotland are between 50 and 60 million years old, whereas the granites of Cornwall, which have been worked for years for their tin and china clay, are about 400 million years old.

The oldest rocks found in the Earth's crust are 3.4×10^9 years old. The Earth must therefore be at least as old. In fact, the Earth is thought to be considerably older than this; it probably originated 4.6×10^9 years ago. If the oldest rocks are 3.4×10^9 years there is obviously a gap of 1.2×10^9 years. We shall discuss this difference in Units 26 and 27.

As we said earlier, the fossil remains of animals and plants are to be found in various rocks. Having measured the age of rocks by radiometric means, one can determine the age of the fossils. Some of the very oldest rocks contain the organic chemical substances,† amino-acids—the building bricks of life (you will learn more of these later in Units 10 and 14). But fairly complicated animals first appeared 0.6×10^9 years ago (some of these are still to be found on present sea-shores virtually identical to their ancient ancestors). Compare this period to man's span on Earth—a mere one million years during which time he has evolved from a recognizable though rather hairy ancestor!

If you wish you may now turn to Appendix 6 (Black) on exponential functions. This pursues the topic of an exponential function a little further; it introduces you to the idea of a *mean-life* and also to the base, 'e'.

2.5.5 Measurement of short time intervals

Finally we turn to the measurement of short time intervals. You saw in our discussion of photography that one way to see the details of fast-changing processes was to use high-speed cinematography. In this technique, the camera shutter is kept open and the film moves continuously. The experiment or process to be recorded is periodically illuminated by intense flashes of light of very short duration. These flashes are repeated at regular intervals and the successive images are formed on the film alongside each other. Figures 37 and 38 show two examples of rapidly changing phenomena photographed by short-duration flashes. The explosion of Figure 39 provided its own intense illumination, so in this case the technique was to use a fast shutter to expose the film for only a short time.

Counting the number of pictures taken between the beginning and the end of a sequence, and knowing the rate at which the pictures are taken, one can calculate the time interval that has elapsed. Modern high-speed cameras can take pictures at the rate of several million per second; by this method therefore it is possible to resolve time intervals of a fraction of a millionth of a second.

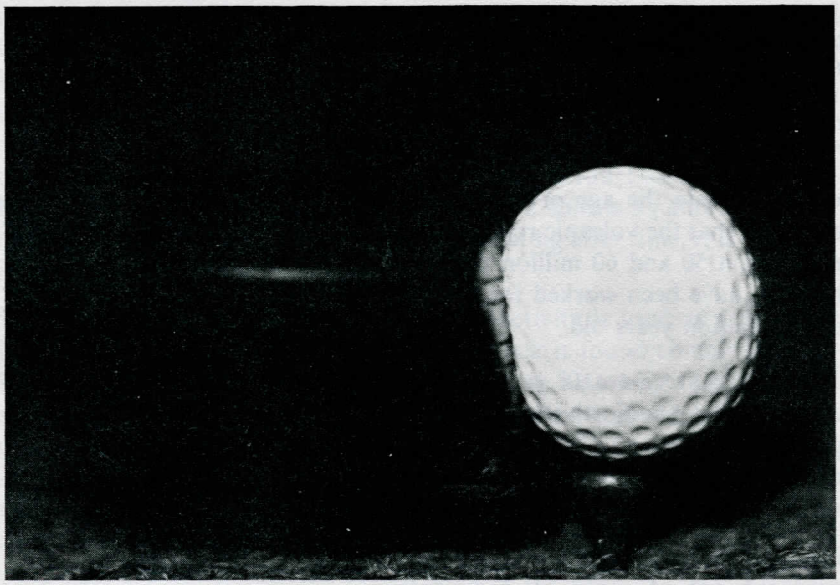


Figure 37

A golf ball deformed at the moment of impact.

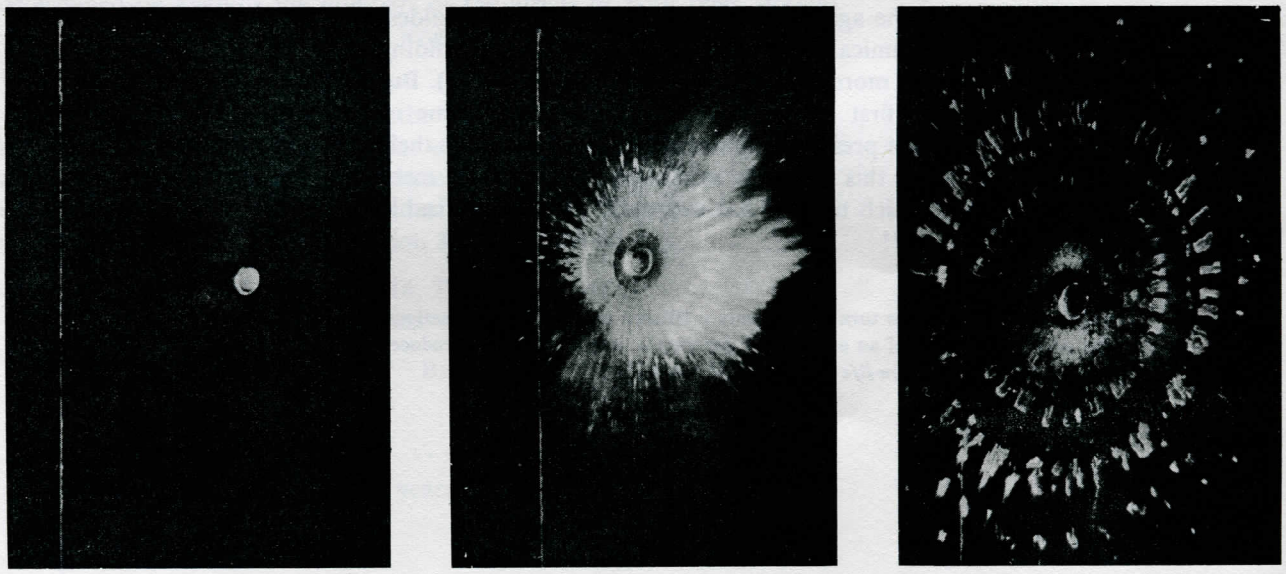


Figure 38

A bullet striking a steel block. The lead in the bullet apparently liquefies, splatters, and then resolidifies.

The *cathode-ray oscilloscope* provides another approach to the problem of measuring short time intervals. This instrument is very versatile and is widely used in many laboratories. In order to understand its detailed construction one has to have a knowledge of electronic circuits. Nevertheless it is not difficult to appreciate the basic principles of its operation.

In Figure 40 we show a simplified diagram of its construction. On the left is a wire filament through which a current is passed. This heats a metal plate called a cathode, which emits the electrons that are to be accelerated. The rate at which the electrons are accepted for acceleration could be modified by making changes to the temperature of the cathode, but in practice it is more convenient to have a metal grid with a hole in it called a control grid. Depending on how one charges the grid relative to the cathode one is able to vary the electric field in the immediate vicinity of the cathode. This in turn affects the number of electrons passing through the hole in the grid.

The electrons passing through the grid tend to be moving in various directions; the purpose of the next component therefore is to focus the beam. Acceleration is produced by the accelerating anode (the word 'anode' meaning that it is electrically positive with respect to some other component—the 'cathode').

The electrons on emerging through a hole in this anode pass through two pairs of parallel plates XX′ and YY′, each pair being at right-angles to the other. If an electric field is established between X and X′, the beam will be deviated horizontally, the direction and magnitude of the deflection depending on the direction and magnitude of the field. Likewise, an electric field between plates Y and Y′ produces vertical deviation. Therefore, a suitable combination of electric fields across XX′ and YY′ directs the beam to any desired point on the screen.

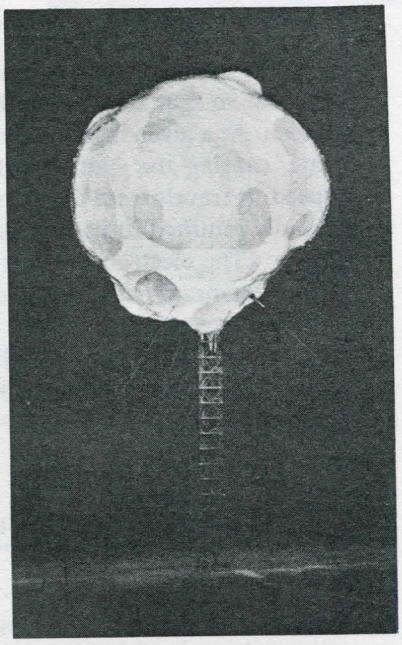


Figure 39

A nuclear bomb explosion at an early stage.

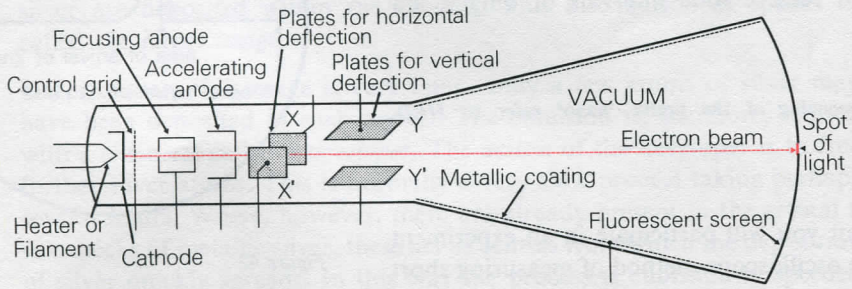


Figure 40

Simplified diagram of a cathode-ray tube.

On reaching the screen, the electrons cause the coating on the inside of the screen to emit visible light. The principle, as you can see, is very similar to that of the electron microscope. The brightness of the spot depends on the speed of the electrons and also on the number of electrons arriving at that point.

What controls this number?

This number can be adjusted by altering the field between the control grid and the cathode.

What component controls the size of the spot?

The size of the spot is governed by the focusing anode.

For the measurement of time intervals, the field between X and X′ is first set so that the spot of light is at the extreme left-hand edge of the screen. At a given instant, the field across XX′ begins to change at a steady rate, causing the spot of light to move across the screen at a constant speed. It travels from the extreme left to the extreme right. The field is then rapidly returned to its original value, taking the spot to the left-hand side again. (The electron beam is suppressed so that its 'return' is not seen.) This horizontal path for the spot of light is called the *time-base*. The time interval over which the field on XX′ is allowed to change is the same as that taken for the spot of light to perform one complete sweep across the screen. By inscribing a scale on the screen one may subdivide the length of the time-base.

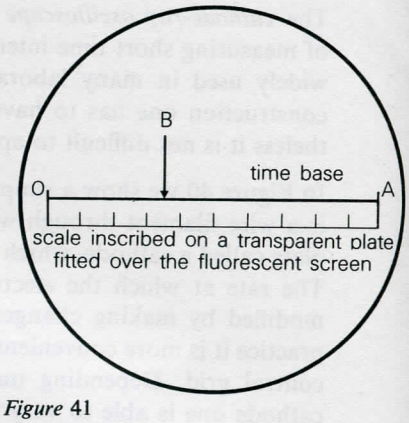


Figure 41

In Figure 41, if it takes the spot 10^{-6} seconds to traverse the complete time-base, OA, estimate how long it takes to reach the point B.

OB reads 3.5 divisions and OA 10 divisions. ∴ The time for the spot to reach B is $(3.5/10) \times 10^{-6} = 0.35 \times 10^{-6}$s

In order to measure the time taken to complete a given process, it is necessary to arrange for two electrical signals to be sent to the oscilloscope —one at the time the process is initiated and another when it is completed. The first pulse signals to the oscilloscope that it should start sweeping the spot across the face of the screen (we say it 'triggers the time base'). The other is applied to one of the plates of the YY′ pair so as to produce a sudden change in electric field in the vertical direction. Because electrons have such a small mass, their motion is easily altered by forces acting upon them. So the electron beam rapidly responds to the force exerted on it by the change in the vertical field, and the spot of light on the screen deviates from its hitherto steady horizontal motion. In Figure 42 you see a typical trace. Clearly by measurement of the distance OB one can calculate the time between the two signals.

By this method it is possible to resolve time intervals of only a few nanoseconds.

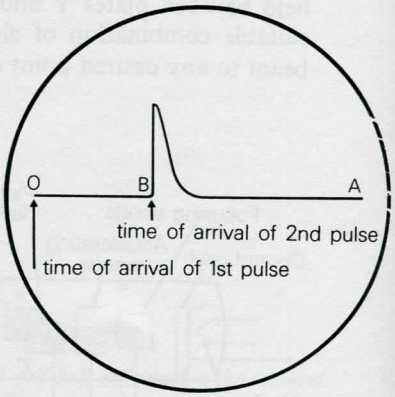


Figure 42

If you have forgotten the meaning of the prefix 'nano' refer to *HED*, section 3.1.5.

In the TV component of this Unit you will participate in an experiment in which you will make use of the oscilloscope method of measuring short time intervals.

Your TV set is very similar in construction to an oscilloscope and to a limited extent can be used as one.

Try home-kit experiment (Unit 2, Experiment 3).

This concludes our brief survey of some of the ways in which measurements of length and time are being extended. In the next Unit you will be studying more closely what exactly one means by the terms 'length' and 'time'.

Appendix 1 (White)

A Short Description of the Principles underlying the Photographic Technique

Certain substances react to light—that is why fruits ripen and fabrics fade in sunshine. Silver bromide is the name of a chemical compound that responds rapidly to exposure to electromagnetic radiation. A compound is a substance that is made of two or more atoms combined in fixed proportions. Silver bromide consists of atoms of silver and of bromine combined in the ratio one-to-one.

When atoms of different types join together, the physical properties of the resulting compounds may differ markedly from those of the separated constituent atoms. Silver bromide is a light-coloured crystalline substance, whereas silver on its own is, of course, a shiny opaque metal, and bromine is a red-brown liquid.

The effect of electromagnetic radiation falling on a silver bromide crystal* is to cause a few electrons to leave their parent atoms. They are then believed to bounce around inside the crystal until they eventually become trapped, usually at the crystal's surface. There they congregate and, because of their electric charge, succeed in disrupting certain of the bonds binding the silver and bromine atoms. The silver atoms, on being released from their bromine atoms, revert to the normal behaviour characteristic of silver, i.e. they become specks of opaque metal.

In the photographic technique, a great many of these tiny crystals are embedded in gelatine and this so-called 'emulsion' is coated on to a strip of cellulose acetate (or on to a glass plate) which acts as a supporting base. Wherever light falls on this photographic emulsion tiny specks of silver are deposited within the crystals. This distribution of specks is called the *latent image*.

Unfortunately this image is very faint—only a few atoms of silver may have been deposited in each crystal. The emulsion is, therefore, treated with a chemical called a *developer*. The action of the developer is to free further silver atoms. This is normally a very slow process taking perhaps several hours. Where, however, there are already present in the crystal a few specks of metallic silver, these act as centres from which the deposition of silver quickly spreads. In this way the process is enormously speeded up and it takes only a matter of a few minutes to develop the whole crystal, i.e. deposit its silver atoms as metal.

Figure 43 shows a highly magnified ($\times 2500$) picture of an emulsion that is only partly developed. You can see quite clearly the metallic silver being deposited around centres located on the surfaces of the crystals.

The process is rather analogous to the onset of an avalanche. Initially a single small stone might be dislodged and this of itself might seem insignificant. If, however, the conditions are right, it can quickly lead to whole rocks being dislodged. In the photographic process it has been calculated that, for each initial metallic silver atom present in the latent image, there are 10^9 after development.

The period for which the emulsion is allowed to remain in the developer

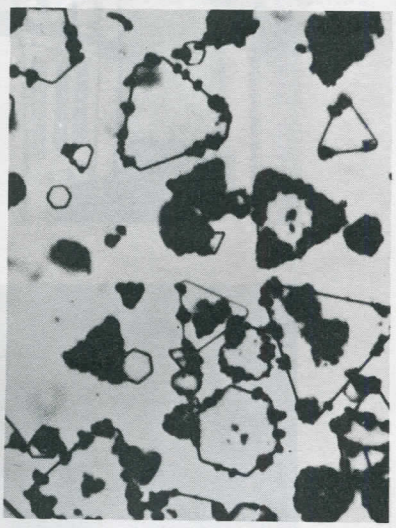


Figure 43

Silver bromide crystals at an early stage of development showing metallic silver being deposited at the crystal surfaces.

* *A crystal is a quantity of substance consisting of atoms arranged in a regular pattern in space. Though there exist liquid crystals, most are solid bodies of regular shape bounded by plane surfaces. Most solids are composed of crystals.*

is chosen so that most of the crystals that were exposed to the radiation are developed, and only a few of those that were unexposed.

After development the emulsion is immersed in another chemical called a *fixer*.

What would you guess was the purpose of this?

It removes the residual silver bromide. This is necessary, otherwise subsequent exposure to light would cause further darkening.

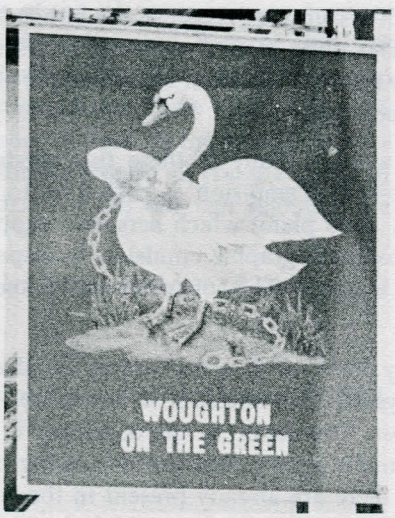
The process we have described produces a 'negative', i.e. a picture where all the light areas show up dark and *vice versa*. To produce a 'print' from the negative one has to reverse the light and dark areas once more.

This is done by shining light through the negative and onto a sheet of paper coated with emulsion. This in turn is developed and becomes the final picture.

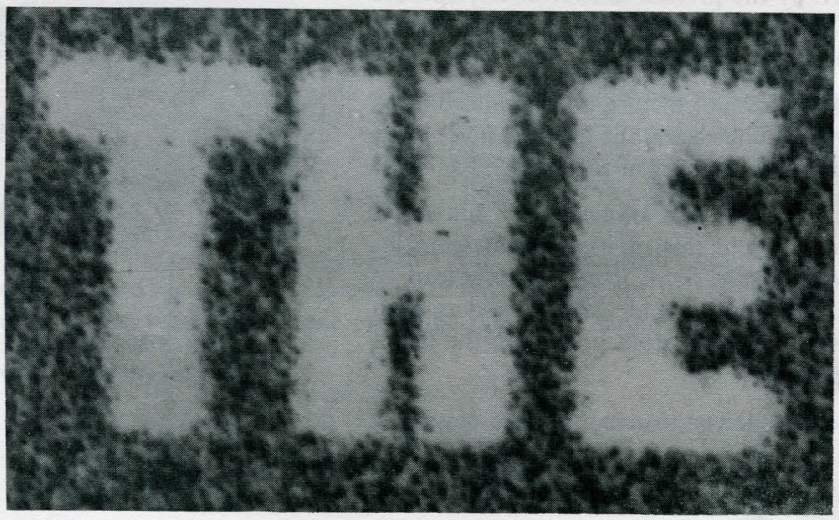
If the photograph of Figure 44 is looked at under progressively greater magnifications, the grainy structure of the picture becomes obvious (Figs. 44 (b) and (c)).

We have spoken only of silver bromide; there are indeed other chemicals that are sensitive to electromagnetic radiation. Various of these chemicals or combinations of them can be made to give emulsions sensitive to radiation over different ranges of wavelength.

(*a*)



(*b*)



(*c*)

Figures 44 (*a*), (*b*), *and* (*c*)—*a photograph seen under successively greater magnification; showing its grainy structure.*

Appendix 2 (White)

The Principle of the Microscope

The microscope is a device for enabling one to see magnified images. In its simplest form it consists of a single lens. This is called a simple microscope or more commonly a *magnifying glass*.

The closer an object is to the eye the bigger it looks. But if it is brought too close, the lens of the eye cannot refract the rays sufficiently to focus them on the retina.* A lens placed in front of the eye introduces additional bending of the rays; this enables focusing to be achieved even with the object quite close (Fig. 45).

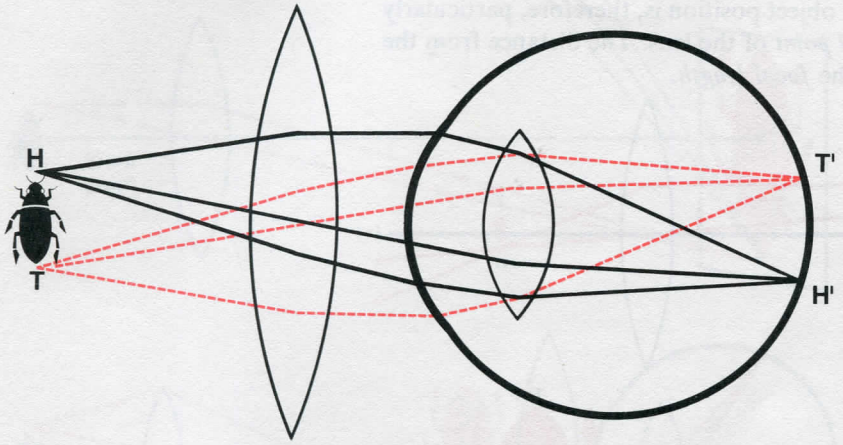


Figure 45

Diagram showing how a magnifying glass supplements the focusing properties of the eye itself.

You are supplied in your home experimental kit with a lens. Use it as a magnifying glass. Note that the image you see through it is the 'right-way-up' —in spite of the image on the retina being inverted (Figure 45). If you do not understand why this is so refer to Appendix 5.

Auxiliary lenses are, of course, also used in spectacles to correct defective vision.

The magnifying glass has its limitations. For one thing there is a limit as to how close the object can be brought up to the eye, due simply to the physical dimensions of the auxiliary lens.

At a further level of sophistication, one has the *compound microscope* (normally just referred to as 'the microscope'). The principle of the microscope is that a lens called the *objective lens* produces a magnified image of the object. The magnifying glass, instead of being used to look at the object directly, now views this already magnified image.

* *This situation is shown in Figure* 60, *Appendix* 5, *p.* 56.

Figures 46 (a), (b), (c), (d), show what happens to the rays passing through a lens when the object is brought up closer to it.

Roughly, how do you expect the light rays to be for Figure 46d?

For the answer turn to Figure 46 (e), at the foot of p. 51, where the rays are drawn.

You see from this series of figures that the closer one brings the object, the further away the image is formed and the bigger it becomes. In Figure 46 b, by the similar triangles HOL and H′O′L (see *MAFS*, section 2.E, if necessary):

$$\frac{\textit{Size of image}}{\textit{Size of object}} = \frac{\textit{distance of image to lens}}{\textit{distance of object to lens}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

A stage is reached however (Fig. 46 c) when the lens is no longer capable of making the rays converge. This object position is, therefore, particularly noteworthy and is called the *focal point* of the lens. The distance from the focal point to the lens is called the *focal length*.

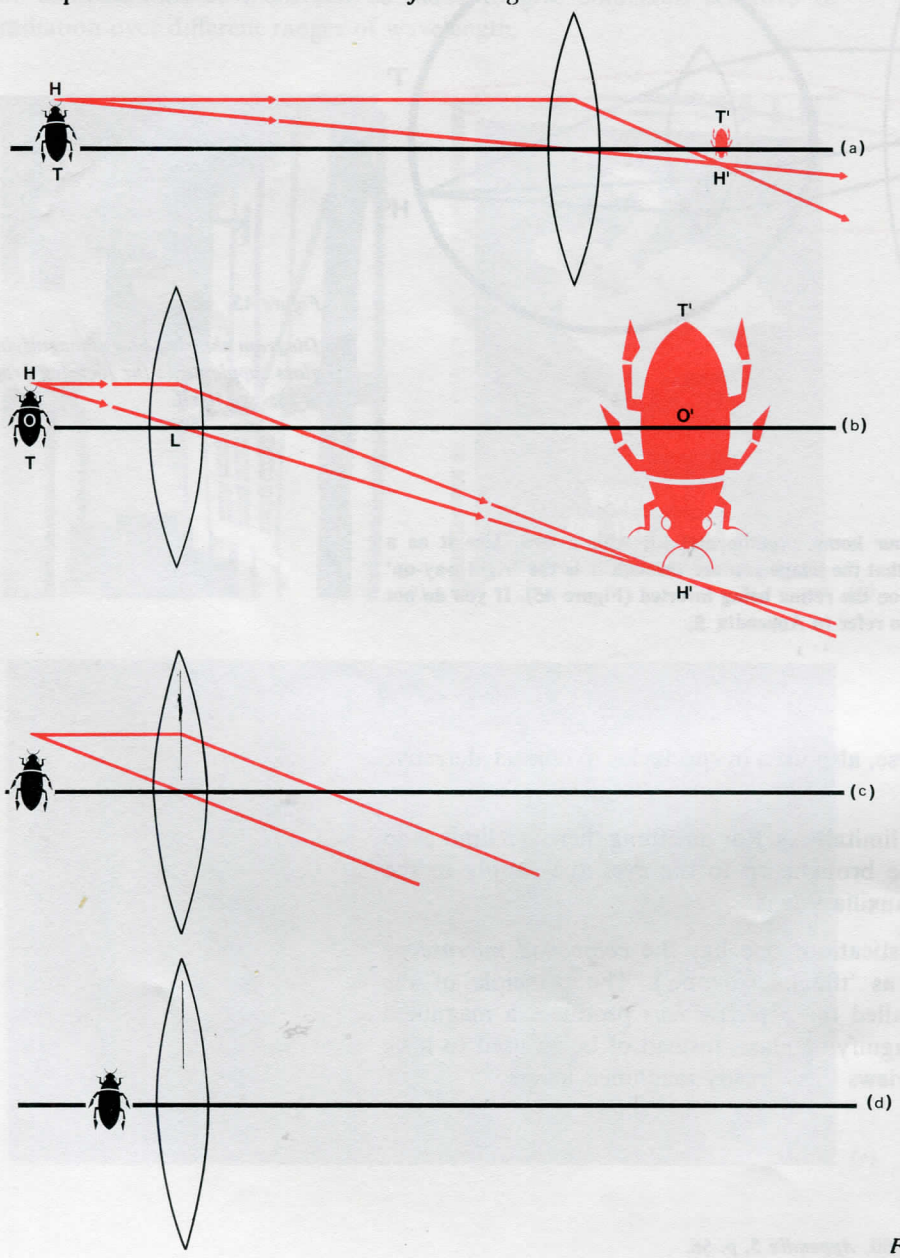


Figure 46

So you see that if the object is placed just a little further back from the lens than its focal point one obtains a large *inverted image*. To see this image one has only to place a screen at the position of the image (Fig. 47).

Alternatively one can look at it with a magnifying glass as in Figure 48. Here one has, in fact, combined the arrangements given in Figures 45 and 46 b. The lens nearest the object is the *objective lens* and the other (our original 'magnifying glass') the *eye-piece lens*.

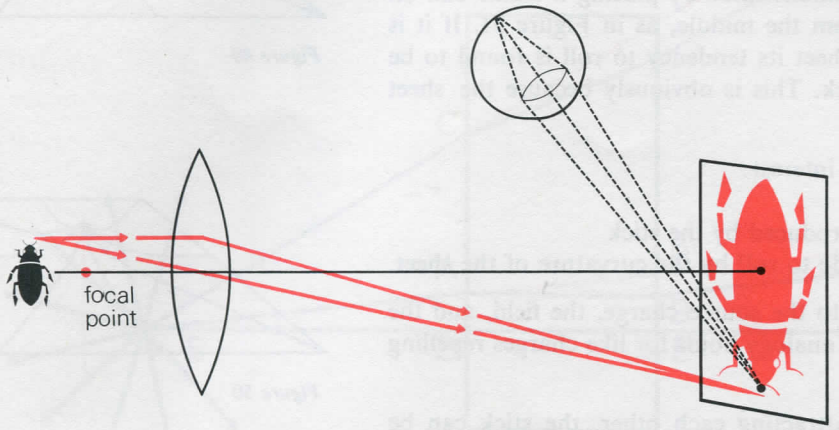


Figure 47

Diagram showing that an object placed a little further from the lens than the focal point produces a large inverted image.

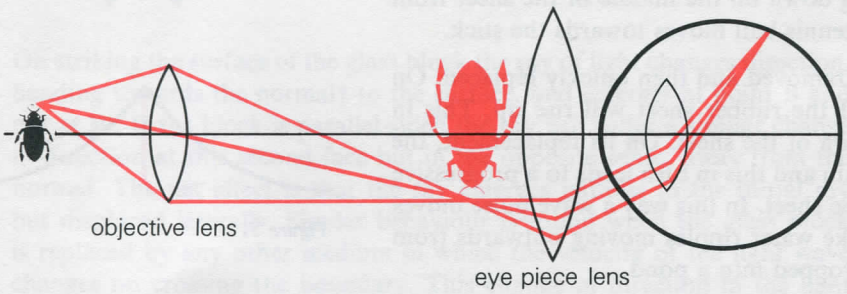


Figure 48

The optics of the microscope.

Figure 48 shows the paths of the rays of light as they pass through the microscope and into the eye.

From equation 6 you see that in order to increase the size of the final image one can either increase the image distance or decrease the object distance to the lens. The first alternative is not of much help. If the microscope is to be at all convenient to handle it cannot have an enormously extended length. The second alternative is the more promising—the object is brought up closer to the objective lens.

Are there any restrictions on how close it can be brought?

Recalling that the object must be kept further away from the lens than the focal point it is clear that in order to reach the very highest magnifications it is necessary to have very thick lenses with short focal lengths. This is why the more expensive microscopes have a selection of objective lenses permitting different magnifications.

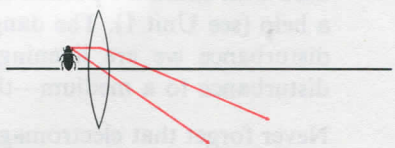


Figure 46 (*e*)

Answer to question posed at head of *p.* 50.

Fields

Let us see if the following analogy can help you get a better feel for this concept of a *field*.

Imagine a very large rubber sheet stretched out tight as in Figure 49. Initially it is flat. Then the middle of the sheet is pushed up with a stick as in Figure 50. The stick is seen to affect the whole sheet—the sheet is now in a different 'condition' to what it was when it was flat.

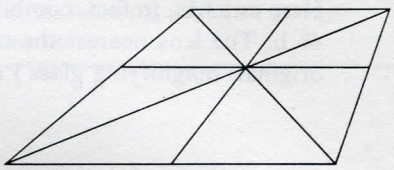


Figure 49

This changed condition can be demonstrated by placing a tennis ball on the sheet. The ball rolls away from the middle, as in Figure 51. If it is placed at various points on the sheet its tendency to roll is found to be greater the closer it is to the stick. This is obviously because the sheet is steeper close to it.

Note there are three quantities of interest:

(1) the stick
(2) the curvature of the sheet produced by the stick
(3) the tennis ball which is made to roll by the curvature of the sheet.

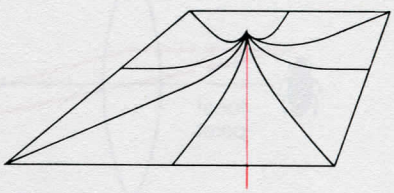


Figure 50

These are analogous respectively to the source charge, the field, and the body acted upon by the field. The analogy holds for like charges repelling each other.

For the case of unlike charges attracting each other, the stick can be replaced by a second one pressing down on the middle of the sheet from above, as in Figure 52. Now the tennis ball moves towards the stick.

Suppose the stick is momentarily removed and then quickly replaced. On its removal, the centre portion of the rubber sheet will rise up. This in turn affects an ever-widening region of the sheet. On its replacement, the centre portion is pushed down again and this in turn leads to a progressive lowering of the other regions of the sheet. In this way a wave pulse moves outwards over the sheet, rather like water ripples moving outwards from a point where a stone has been dropped into a pond.

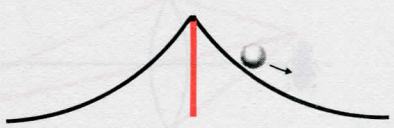


Figure 51

Note that after the stick has been replaced one cannot, from a study of the stick alone, see all that is going on—one cannot predict that at some subsequent time a distant tennis ball lying on the sheet is going to receive a jolt. The behaviour of this ball can only be explained by reference to the travelling distortion on the rubber sheet.

This is analogous to our source charge being moved from one position to another and back again; to understand the subsequent movement of a distant charge one must concentrate attention on the field.

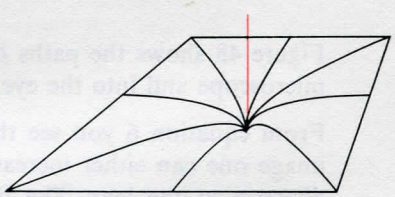


Figure 52

Warning

Analogies are all very well and this one may serve to give you a helpful mental picture for visualizing these fields. However, analogies must be used with caution—pushed too far they become more of a hindrance than a help (see Unit 1). The danger with this particular analogy is clear. The disturbance we are likening our electromagnetic wave to is in fact a disturbance to a medium—the rubber sheet. No rubber sheet—no wave.

Never forget that electromagnetic waves do not require a medium.

Appendix 4 (Red)

Lenses

In Figure 53, light is diverging in all directions from a source S. Some will pass from the source to a given point A on the surface of a glass block. SA is one of the many possible paths that light emitted from S can take—it is called a *ray*.

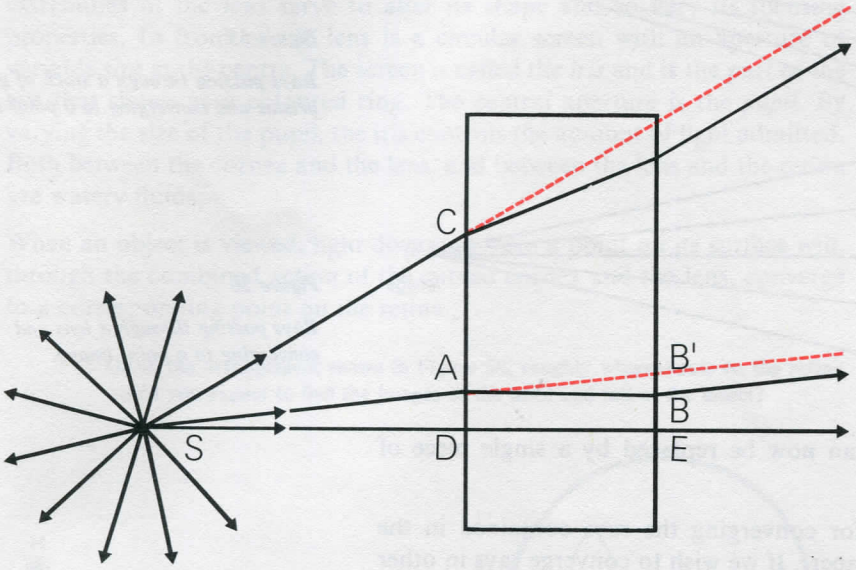


Figure 53

Diagram showing rays of light from source, S, being refracted on passing through a parallel-sided glass block.

On striking the surface of the glass block the ray of light changes direction, bending towards the normal† to the surface, and emerges at point B and not at B′. If the block is parallel-sided the ray will suffer a similar change in direction at this second face but in the opposite sense, away from the normal. The net effect is that the ray emerges parallel to the initial ray but displaced laterally. Similar behaviour is found when the glass block is replaced by any other medium in which the velocity of the light wave changes on crossing the boundary. This change of direction in the light when passing from one medium into another is called *refraction*.

The amount of bending depends upon the angle at which the ray strikes the surface. This is illustrated by rays SC and SD. In the latter case, where it meets the surface at right-angles, i.e. normally, it passes on undeviated and emerges at E. If the block were not parallel-sided, the emergent rays in general would not only be laterally displaced but also show a resultant change in direction as in Figure 54.

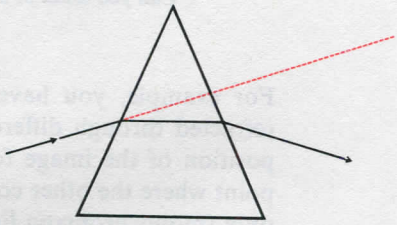


Figure 54

Diagram showing a ray of light emerging from a wedge-shaped glass block and having changed its direction.

What factors do you think govern the magnitude of this change of direction?

This overall change in direction will depend on the angle at which the ray strikes the first face, the angle between the faces, on the exact nature of the glass and on the wavelength of the light. Now imagine a set of glass prisms (i.e. wedge-shaped blocks of glass) stacked on top of each other, each with a slightly different angle between its opposite faces compared to that of its neighbour. These angles are arranged so that all the light rays coming from the source and passing through the stack reconverge to a single point, I, called the *image* (Fig. 55).

By making the individual prisms smaller and correspondingly increasing their number, the surfaces of the stack become smooth, as shown in Figure 56.

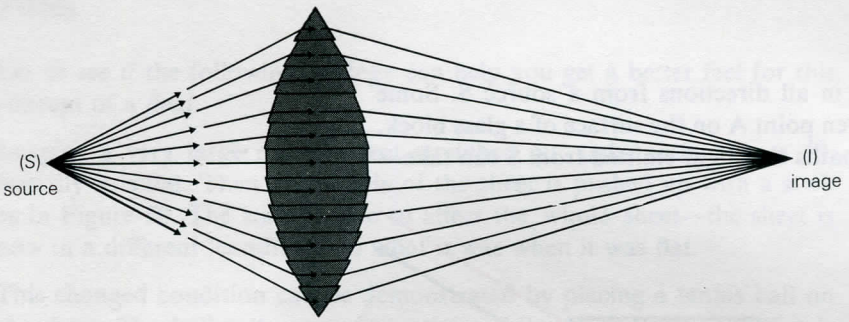


Figure 55

Rays passing through a stack of glass prisms and converging to a point image.

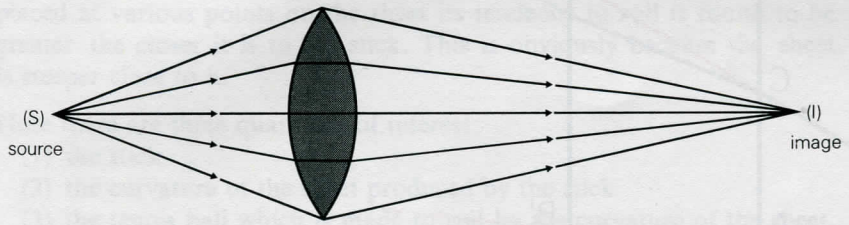


Figure 56

Rays passing through a lens and converging to a point image.

The stack of separate prisms can now be replaced by a single piece of glass called a *lens*.

This arrangement is all right for converging the rays contained in the plane shown (the plane of the paper). If we wish to converge rays in other planes as well then the lens must have the same profile in all planes—in fact it is found that both surfaces have to approximate to portions of *spheres*.

This is a very simple picture of the way in which a spherical lens forms images. In practice there are departures from this idealized situation; the rays do not all converge to a single point and the resulting defects are called *aberrations*.

Can you think of any reason why such defects are to be expected?

For example, you have just read how light of differing wavelength is refracted through different angles. If point I in Figure 56 is the correct position of the image for one component of the light it will not be the point where the other colours will converge. Even with light of one colour only (monochromatic light) there are a variety of aberrations depending on the position of S relative to the lens, and also on the size of the lens.

As lenses are an integral part of many optical instruments, for example the microscope and telescope, the reduction of aberrations to a minimum is one of the chief problems in their design.

Appendix 5 (Red)

The Optics of the Human Eye

The human eye is essentially an optical instrument. Light enters through the transparent *cornea* (Fig. 57) and in so doing is refracted. Its direction is further modified by the *lens*. The light is focused on to the sensitive inner layer of the rear of the eye called the *retina*. Muscles attached to the extremities of the lens serve to alter its shape and so vary its focusing properties. In front of the lens is a circular screen with an aperture of variable size at the centre. The screen is called the *iris* and is the part of the eye that shows as a coloured ring. The central aperture is the *pupil*. By varying the size of the pupil, the iris controls the amount of light admitted. Both between the cornea and the lens, and between the lens and the retina are watery fluids.

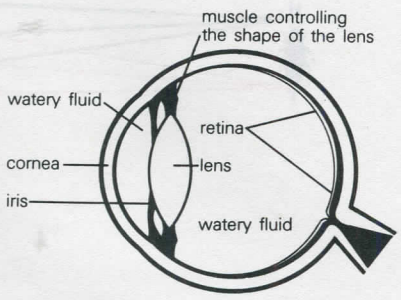


Figure 57

The human eye.

When an object is viewed, light diverging from a point on its surface will, through the combined action of the curved cornea and the lens, converge to a corresponding point on the retina.

Given the arrangement shown in Figure 58, roughly whereabouts on the retina would you expect to find the images of the head and tail of the insect?

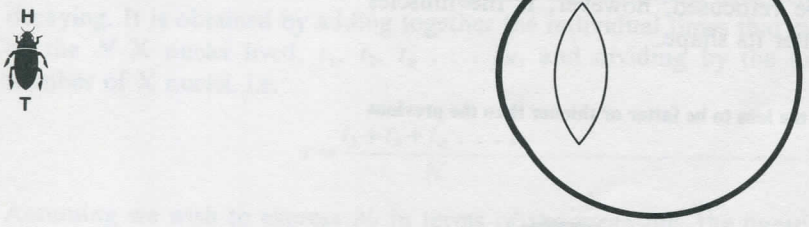


Figure 58

Along the central axis of the eye, the two faces of the lens and the surface of the cornea are parallel. When a ray passes through parallel faces you have already seen that it goes on undeviated (it will be slightly displaced laterally, but you need not worry unduly about that). So if one were to draw a ray passing from a point on the insect's head, H, and through a point somewhere between the cornea and the back face of the lens, it would carry on to the retina undeviated. This serves to locate the position of the image of the head, H′ (Fig. 59). The same procedure serves to locate the image T′ of the tail, T. As you can see the image turns out to be inverted (the reason one mentally 'sees' the image the right way up is to do with the way the brain interprets the signals coming from the retina).

If the shape of the lens is such that the object is correctly focused on the retina one can now draw in any rays one pleases. We show in Figure 59 two further rays coming from the head and converging on the same point of intersection with the retina as our earlier ray. We also show two additional rays coming from the tail.*

* *You will notice in this diagram that we have drawn the rays as though they change direction inside the lens, this of course, does not happen; refraction takes place at the two* surfaces *of the lens and it is these two changes in direction that give the combined overall deviation. However it is a* convention *when representing rays passing through a lens, to extrapolate the incident and emergent rays to meet as shown. From now on we shall adopt this convention.*

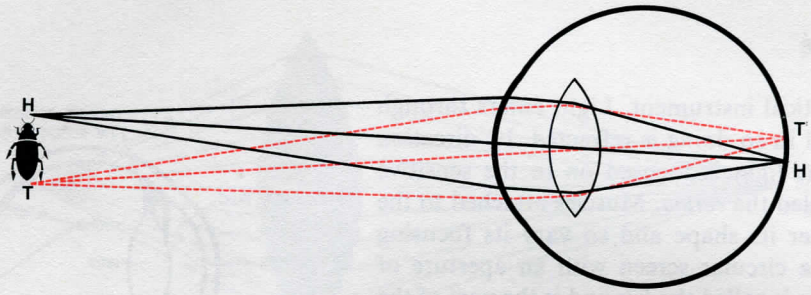


Figure 59

Diagram showing how an image is formed on the retina of the eye.

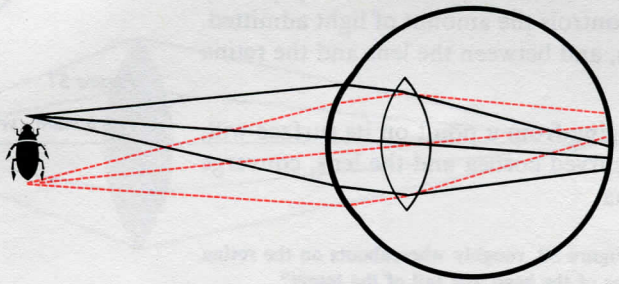


Figure 60

Diagram showing that when an object is brought closer to the eye and the shape of the lens remains unaltered, the image is no longer correctly focused.

If the object is now brought closer to the eye, as in Figure 60, without making any adjustment to the shape of the lens, the rays from each point on the object will not converge properly on the retina—the image will appear blurred. The image can be refocused, however, if the muscles attached to the lens are made to alter its shape.

Do you expect the new shape of the lens to be fatter or thinner than the previous one?

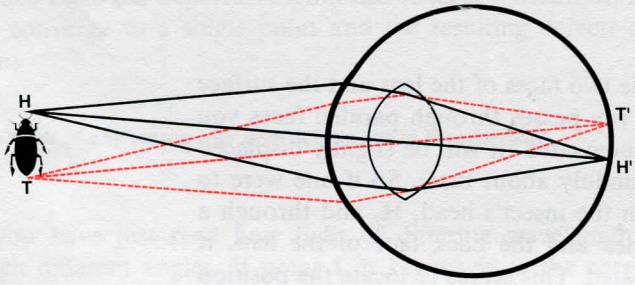


Figure 61

Diagram showing the lens altered in shape in order to focus the image with the object in its new position.

Figure 61, shows the shape of the lens after this refocusing has been done. You will note that in Figure 61 the image is bigger on the retina than in Figure 59 when the object was further away. You might think then that in order to see smaller and smaller objects all one has to do is to bring them up closer.

But what do you find happens when you bring an object up very, very close to your eye? Try it.

There is a limit to the amount by which the lens of the eye can alter its shape. At distances of little less than that at which you are reading this print, your eyes can no longer be made to bend the rays sufficiently to secure good focus—you get the kind of situation you saw earlier in Figure 60. This inability tends to become more marked with advancing age.

Appendix 6 (Black)

Exponential Functions

Look at the working of the answer to the problem following equation 3 (p. 41): you will note that instead of writing $N_t = 1\,600\,(\frac{1}{2})^{12/3}$, we could equally well have written $N_t = 1\,600\,(\frac{1}{4})^{12/6}$ or $1\,600\,(\frac{1}{16})^{12/12}$. This means we are not restricted to expressing the speed of the reaction by a half-life; if we so chose we could express it in terms of 'a quarter-life', i.e. the time for the initial number of nuclei to be reduced to a quarter (6 hours), or as 'a sixteenth-life' (12 hours). In practice, one never uses these particular terms, but they do serve to illustrate something important—there are an infinite number of ways of specifying the rate of the transformation. We can express it in a general way in terms of a time interval T_p, such that a fraction p of the original sample remains at the end of this period. Then

$$N_t = Np^{\,t/T_p} \qquad \text{.....(7)}$$

Make sure that you understand what each of these symbols stands for. If you have forgotten any, refer back to p. 41, bearing in mind that there p was taken to be $\frac{1}{2}$.

It is customary to use the half-life and also one other value of T_p. This second value is called the *mean-life*, τ (a Greek letter pronounced 'taw'). The mean life for a reaction is the average time that a nucleus lives before decaying. It is obtained by adding together the individual times that each of the N X nuclei lived, t_1, t_2, t_3 . . . t_N, and dividing by the total number of X nuclei, i.e.

$$\tau = \frac{t_1 + t_2 + t_3 \ldots t_N}{N} \qquad \text{.....(8)}$$

Assuming we wish to express N_t in terms of the mean-life, the question naturally arises as to what value does p assume in equation (7) for such a value of T_p, i.e. what fraction of the X nuclei remain after time, τ? Taking the mathematical form of the curve in Figure 36 (p. 41) and using calculus one finds the answer to be

$$p = \frac{1}{2.718\ldots}$$

The denominator is an example of an irrational number (it continues to an infinite number of decimal places). It is a number that appears in a great many scientific problems and because of its importance it is assigned a special symbol, e (this is similar to the way we assign the special symbol, π, to another irrational number, 3.1415 . . ., calculated from the ratio of the circumference to the diameter of a circle).

$$\text{Thus,} \qquad N_t = N\left(\frac{1}{e}\right)^{t/\tau}$$

$$\text{or inverting,} \qquad N_t = N\,e^{-t/\tau} \qquad \text{.....(9)}$$

The fraction of nuclei remaining after time, τ, i.e. $\left(\frac{1}{e}\right)$, is 0.368. Thus, fewer than half of the nuclei remain; the mean-life is therefore greater than the half-life. How much greater?

Rearranging equation (9):

$$\frac{N}{N_t} = e^{t/\tau}$$

At time $t=T_{\frac{1}{2}}$,

$$N_t=\tfrac{1}{2}N$$

Thus $N/\tfrac{1}{2}N=e^{T_{\frac{1}{2}}/\tau}$

i.e. $2=e^{T_{\frac{1}{2}}/\tau}$

Taking logarithms of both sides

$$\log_{10}2=(T_{\frac{1}{2}}/\tau)\log_{10}e$$
$$=(T_{\frac{1}{2}}/\tau)\log_{10}2.718$$
$$0.3010=0.4343T_{\frac{1}{2}}\tau$$
$$T_{\frac{1}{2}}=\frac{0.3010}{0.4343}\tau$$
$$T_{\frac{1}{2}}=0.693\tau \quad \text{.....(10)}$$

This gives us the relationship between the half-life, $T_{\frac{1}{2}}$, and the mean-life, τ.

Normally the speed of a radioactive decay is characterized by its half-life. In the decays of sub-nuclear particles however it is customary to use the mean-life. The mean-life is particularly useful as its inverse is the *decay constant*, λ, of the reaction,

$$\frac{1}{\tau}=\lambda \quad \text{.....(11)}$$

The decay constant is the constant of proportionality relating the number of nuclei decaying per unit time to the number of nuclei present, i.e.

$$-\frac{dN}{dt}=\lambda N$$

or $$-\frac{dN}{N}=\lambda dt \quad \text{.....(12)}$$

Where dN denotes 'a small change in N', and dt is 'a small change in t'. (The letter 'd' stands for 'differential'.) The negative sign indicates the change in N is a decrease.

If you look at the answer to the problem following equation (5) (p. 43) you see that there is only a comparatively small change in N (just over 2 per cent, so you should be able immediately to work out the decay constant for the reaction:

$$^{87}Rb \rightarrow {}^{87}Sr$$

Try it.

$$-\frac{dN}{N}=\frac{0.015}{0.680+0.015}$$
$$\approx 0.022$$
$$dt=1.5\times10^{9}\text{ years}$$
$$\therefore\ \lambda\approx\frac{0.022}{1.5\times10^{9}}\simeq 1.5\times10^{-11}\text{/year}$$

From this value of λ now obtain a value for the mean life-time, τ:

$$\tau=\frac{1}{\lambda}$$
$$\tau=6.7\times10^{10}\text{ years}$$

From the relation between $T_{\frac{1}{2}}$ and τ given earlier, equation (10), we can check our result for τ

$$T_{\frac{1}{2}}=0.693\,\tau$$
$$\therefore\ \tau=\frac{4.7\times10^{10}}{0.693}=6.8\times10^{10}\text{ years}$$

This agrees within the accuracy of our calculations.

Appendix 7 (White)

Glossary

CYCLE Any series of changes or operations performed by or on a system which brings it back to its original state. E.g., a vibrating object executes one cycle when it travels from one extreme position to the opposite and back again.

GREAT CIRCLE Circle obtained by cutting a sphere by a plane passing through the centre. e.g., regarding the Earth as a sphere, the equator is a great circle, as are all meridians of longitude.

INTEGRAL MULTIPLE A multiplying factor consisting of a whole number i.e. integer.

INTENSITY Generally means amount. e.g., 'intensity of electron beam' means the amount or number of electrons contained in the beam.

MEAN Generally understood to be the arithmetic mean, i.e. the average.

NORMAL A line that is perpendicular to a surface.

ORGANIC CHEMICAL SUBSTANCES Chemical substances containing carbon combined with hydrogen and often other elements.

STANDARD ATMOSPHERIC PRESSURE A unit of pressure. The pressure that will support a column of mercury 760 mm high at 0° C, sea-level and latitude 45°. (Atmospheric pressure fluctuates about this value from day to day.)

SUBTEND Two points, A and B, are said to subtend the angle ACB at the point C.

Sections 2.1 and 2.2

	(a) true	(b) false
Question 1 (*Objective 1*) The wavelength of a wave is the reciprocal of the frequency	☐	☑
Question 2 (*Objective 1*) The velocity of a wave is the product of the frequency and the wavelength	☑	☐
Question 3 (*Objective 1*) Ultrasonic waves are sound waves that are too weak to be heard	☐	☑
Question 4 (*Objective 1*) A sound wave with frequency 30 000 Hz is an example of an ultrasonic wave	☑	☐
Question 5 (*Objective 1*) The amplitude of a sound wave is the maximum speed with which the particles of the medium move as they oscillate about their mean position	☐	☑

Question 6 (*Objective 14*)

When a wave moves into a medium where its velocity is lower,

	(a) increases	(b) decreases	(c) remains unchanged
(i) the wavelength	☐	☑	☐
(ii) the frequency	☐	☐	☑

Section 2.3

Question 7 (*Objective 4*)

Which type of radiation would be used to investigate each of the problems listed below. Tick the appropriate box.

(a) —ultrasonic
(b) —infra-red
(c) —ultra-violet light
(d) —X-rays

	(a)	(b)	(c)	(d)
(i) A document suspected of being fraudulently altered	☐	☐	☑	☐
(ii) The whereabouts of a shoal of fish	☑	☐	☐	☐
(iii) Whether a bone has been fractured	☐	☐	☐	☑
(iv) The successive alterations made to a painting	☐	☐	☑	☐
(v) The area over which a crop has become diseased	☐	☑	☐	☐
(vi) The whereabouts of an army convoy at night	☐	☑	☐	☐

Question 8 (*Objective 5*)

Electromagnetic radiation of the following wavelengths (in metres) are classed as what type of radiation listed on the left? Tick the appropriate box.

	(a) 10^{-4}	(b) 5×10^{-7}	(c) 10^{-6}	(d) 3×10^{-7}	(e) 10^{3}
(i) visible	☐	☑	☐	☐	☐
(ii) ultra-violet	☐	☐	☐	☑	☐
(iii) gamma	☐	☐	☐	☐	☐
(iv) X-ray	☐	☐	☐	☐	☐
(v) infra-red	☑	☐	☑	☐	☐
(vi) radio and TV	☐	☐	☐	☐	☑

Question 9 (*Objective 6*)

When two materials are rubbed together the one that has acquired a positive charge has:

(i) gained nuclei ☐
(ii) gained electrons ☐
(iii) lost nuclei ☐
(iv) lost electrons ☑

Question 10 (*Objective 7*)

Tick the appropriate boxes

	(a) ultra-sonic waves	(b) ultra-violet waves
(i) Require a medium	☑	☐
(ii) The source must move in order to produce the wave	☑	☑
(iii) The quantity that is varying in the wave is directed along the ray representing the path of the wave	☑	☐
(iv) Can exhibit diffraction	☑	☑

Question 11 (*Objective 11*)

The electric force on a small charged body, B, produced by another stationary charged body, A, acts

(i) along the direction AB ☑
(ii) at right-angles to AB ☐
(iii) in some other direction ☐

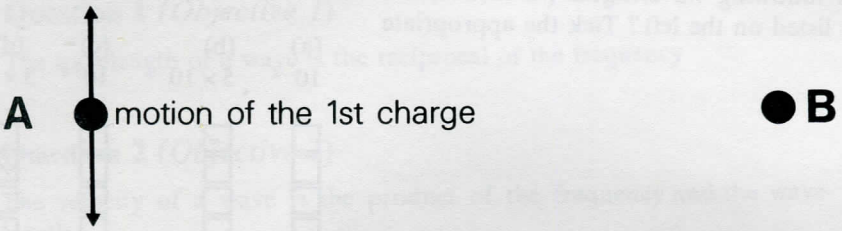
If body A oscillates back and forth in the direction along the line AB, the field at B acts

(iv) along the direction AB ☑
(v) at right-angles to AB ☐
(vi) in some other direction ☐

Question 12 (*Objective 11*)

A positive electric charge, A, oscillates about its mean position as shown in the diagram. Explain briefly what will happen to a second positive charge, B, initially at rest, if this second charge is set free so as to be able

to move in any direction. What effect if any would you expect the presence of the second charge to have on the first charge?



Section 2.4

Question 13 (*Objective 1*)

To obtain a line spectrum light of a single wavelength is needed

(a) true ☐ (b) false ☑

Question 14 (*Objectives 1, 10*)

The standard metre is defined as a length equal to a certain number of wavelengths, measured in air at standard pressure, of a particular reddish-orange line in the spectrum of the gas krypton

(a) true ☐ (b) false ☑

Question 15 (*Objective 9*)

All optical illusions, including after-images, arise from the way the brain organizes the information sent to it

(a) true ☐ (b) false ☑

Section 2.5

Question 16 (*Objective 1*)

A battery is a device for maintaining an electric field around a circuit

(a) true ☑ (b) false ☐

Question 17 (*Objective 1*)

An electron microscope is so-called because it is a microscope for studying electrons

(a) true ☐ (b) false ☑

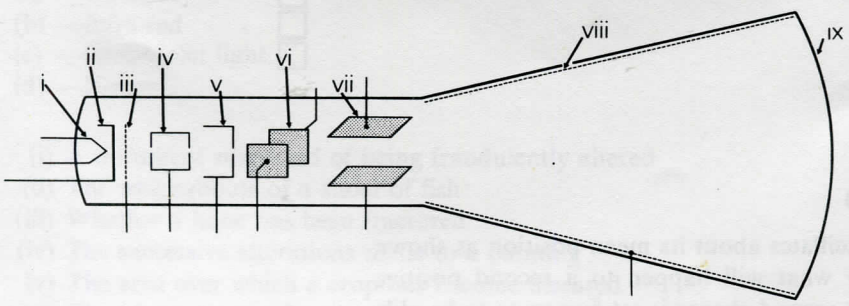
Question 18 (*Objective 1*)

In the triangulation method of distance estimation, the baseline is the distance from the observer to the object being viewed

(a) true ☐ (b) false ☑

Question 19 (*Objective 2*)

Name the parts of the cathode ray tube:



(i)
(ii)
(iii)
(iv)
(v)
(vi)
(vii)
(viii)
(ix)

Question 20 (*Objective 12*)

The following effects all play a part in the production of a beam of free electrons. Place numbers 1 to 7 in the appropriate boxes to indicate the order in which the effects take place

(i) increase in random motion of electrons [4]
(ii) systematic motion of electrons established in a circuit [2]
(iii) electrons focused by focusing anode [6]
(iv) chemical reaction set up in a battery [1]
(v) electrons pass through the accelerating anode [7]
(vi) electrons escape from a metal surface [5]
(vii) electrons collide with a lattice of atoms [3]

Question 21 (*Objective 13*)

The law of conservation of electric charge states that the total amount of positive charge equals the total amount of negative charge

(a) true ☐ (b) false ☑

Question 22 (*Objective 14*)

A tank has a hole in the bottom out of which water is allowed to flow. The height of the water level in the tank drops from 16 cm to 4 cm in 56 seconds. Assuming the rate at which the water leaves is proportional to the height of the water level, calculate how long it would take for the level to drop from 40 cm to 2.5 cm.

Appendix 1

Question 23 (*Objective 8*)

Which items chosen from the right-hand column are associated with those on the left? Indicate your answers below:

(i) Fixer.
(ii) Migration of electrons.
(iii) Emulsion.
(iv) Cellulose acetate.
(v) Developer.

(a) A mixture of silver bromide crystals and gelatine.
(b) Formation of latent image.
(c) Removal of silver bromide.
(d) Deposition of silver on a large scale.
(e) A supporting base.

(i) c
(ii) b
(iii) a
(iv) e
(v) d

Appendix 2

Question 24 (*Objective 1*)

(a) true (b) false

The objective lens of a microscope is the lens closest to the object ☑ ☐

Question 25 (*Objective 14*)

A man six feet tall is photographed at a distance 18 ft from the camera, and his image on the film is found to measure 1 inch.

What is the distance between the lens and the film?

Appendix 5

Question 26 (*Objective 1*)

The iris controls the amount of light admitted into the human eye

(a) true ☑ (b) false ☐

Question 27 (*Objective 3*)

The speed of light in the cornea of the eye is about three-quarters that in air. In the watery fluids the speed is about the same as in the cornea. Its mean value in the lens is somewhat lower. At which boundaries would you expect refraction to take place?

Answer by ticking the appropriate box or boxes.

(i) air—cornea	☐	☐
(ii) cornea—first watery fluid	☐	☐
(iii) first watery fluid—lens	☐	☐
(iv) lens—second watery fluid	☐	☐

The Handling of Experimental Data, Section 3.1.5

Question 28 (*Objective 1*)

One-millionth of 10^{-3} second is a:

Tick correct answer

(i) millisecond	☐
(ii) microsecond	☐
(iii) nanosecond	☑
(iv) none of these	☐

Self-Assessment Answers and Comments

Question 1

Answer (b)

Comment = *v*/*f not* 1/*f*

Question 2

Answer (a) $v = \lambda f$

Question 3

Answer (b)

Comment

Ultrasonic waves cannot be heard because of their high frequency, not because they are necessarily of low intensity.

Question 4

Answer (a)

Question 5

Answer (b)

Comment

The amplitude is the maximum distance a particle of the medium moves from its mean position.

Question 6

Answer (i) b, (ii) c.

Question 7

Answer (i) c, (ii) a, (iii) (d), (iv) c, (v) b, (vi) b.

Comment

With regard to the last item it has been found that the warmth from the engines of the army trucks makes them visible in the infra-red region. Anything that is warm is liable to show up well in the infra-red. You might also have included (iv)d because X-rays can also be used to examine paintings.

Question 8

Answer (v) a, (i) b, (v) c, (ii) d, (vi) e.

Comment

There were no wavelengths specified corresponding to gamma or X-rays.

Question 9

Answer (iv)

Question 10

Answer (i) a, (ii) a and b, (iii) a, (iv) a and b.

Question 11

Answer (i) and (iv)

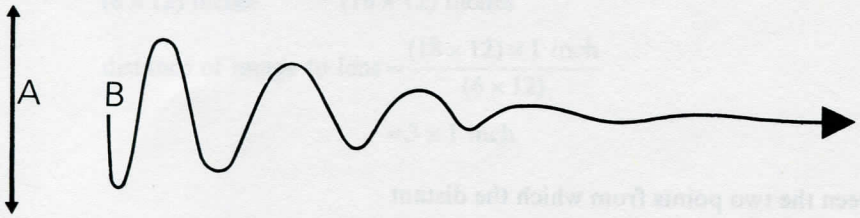
Comment

If you chose (v) instead of (iv) you should note that we are here considering the charge to move *in* the direction of interest, not at right-angles as was the case in the text. Because of this there can be no sideways component of the electric field. The field will always be directed along AB. The effect of varying the distance of A from B will be to vary the strength of the field, not its direction.

Question 12

Answer

B experiences a steady component of force acting directly away from the mean position. It also experiences a fluctuating component at right-angles to this direction. It will therefore move away from A along a path that wanders from side to side as in the diagram.



You might also add that the amplitude of the up and down movements diminishes the further away B gets from A because the strength of the field diminishes with distance. Also the charge moves with ever increasing speed.

Whatever force is exerted by A on B, there will be an equal and opposite force by B on A. Just as B is regarded as being acted upon by the electric field of A, so A is acted upon by the electric field of B. As B moves up and down in response to the up and down movement of A, so A will experience an up and down force due to the movement of B. As B is repelled by A so A is repelled by B.

Question 13

Answer (b)

Comment

Light of a single wavelength is one example of a line spectrum, but generally speaking there is more than one wavelength.

Question 14

Answer (b)

Comment

In vacuum, not in air.

Question 15

Answer (b)

Comment

An after-image is caused by fatigue of the sensory receptors rather than by the manner in which the brain organizes the information sent to it.

Question 16

Answer (a)

Question 17

Answer (b)

Comment

The electron microscope is so called because it *uses* electrons, not because it is designed for studying electrons.

Question 18

Answer (b)

Comment

The baseline is the distance between the two points from which the distant object is viewed.

Question 19

Answer See Figure 40 of main text.

Question 20

Answer (i) 4 (iii) 6 (v) 7 (vii) 3
(ii) 2 (iv) 1 (vi) 5

Question 21

Answer (b)

Comment

The law says that the amount of positive charge less the amount of negative charge must remain constant; it does not say that it must be zero.

Question 22

Answer 112 seconds

4 cm is a quarter of 16 cm, therefore 56 seconds represents twice the 'half-life' for the height of the water level. The half-life is thus 28 seconds. Starting at 40 cm, each successive half-life leaves a level of 20, 10, 5 and 2.5 cm. Therefore to go from 40 to 2.5 cm requires 4 half-lives, i.e. $4 \times 28 =$ 112 seconds.

Question 23

Answer (i) c, (ii) b, (iii) a, (iv) e, (v) d.

Question 24

Answer (a)

Question 25

Answer 3 inches.

$$\frac{\text{size of image}}{\text{size of object}} = \frac{\text{distance of image to lens}}{\text{distance of object to lens}}$$

$$\frac{1 \text{ inch}}{(6 \times 12) \text{ inches}} = \frac{\text{distance of image to lens}}{(18 \times 12) \text{ inches}}$$

$$\text{distance of image to lens} = \frac{(18 \times 12) \times 1 \text{ inch}}{(6 \times 12)}$$

$$= 3 \times 1 \text{ inch}$$

Question 26

Answer (a)

Comment

The amount of light admitted to the eye is governed by the size of the pupil—but this is controlled by the iris.

Question 27

Answer (i), (iii) and (iv).

Comment

Refraction occurs only where there is a change of speed.

Question 28

Answer (iii)

Acknowledgements

Grateful acknowledgement is made to the following sources for illustrations used in this Unit:

P. W. ALLEN AND COMPANY, Fig. 13; CHARLES F. BRANDFORD CO., PUBLISHERS, Newton, Mass., U.S.A., Fig. 38; C. B. DAISH, Fig. 37; DOVER PUBLICATIONS INC., from *Visual Illusions* by M. Luçkiesh, Fig. 17; ELEK BOOKS LTD., from *The New Age in Physics* by Sir Harry Massey, Fig. 34; LABORATOIRE D'OPTIQUE, C.N.R.S., Fig. 33; FRANK W. LANE, Fig. 3; M. J. LANGFORD and THE FOCAL PRESS LTD., Fig. 43; MCGRAW HILL BOOK CO., from *Principles of Optics* by Hardy and Perrin, Fig. 25; NATIONAL AGRICULTURAL ADVISORY SERVICE, E. Region, Ministry of Agriculture, Crown Copyright, Fig. 14; NATIONAL PHYSICAL LABORATORY, Teddington, Crown Copyright, Fig. 20; ST. BARTHOLOMEW'S HOSPITAL AND MEDICAL COLLEGE, Fig. 15; ST. THOMAS'S HOSPITAL MEDICAL SCHOOL, Fig. 12; U.S. ATOMIC ENERGY COMMISSION, Fig. 39; SCIENTIFIC AMERICAN INC., from 'Standards of Measurement' by A. V. Astin, *Scientific American*, June 1968, Copyright © 1968 Scientific American, Inc., Fig. 18.

Notes

S100—SCIENCE FOUNDATION COURSE UNITS

1 Science: Its Origins, Scales and Limitations
2 Observation and Measurement

3 Mass, Length and Time
4 Forces, Fields and Energy

5 The States of Matter

6 Atoms, Elements and Isotopes: Atomic Structure
7 The Electronic Structure of Atoms

8 The Periodic Table and Chemical Bonding
9 Ions in Solution

10 Covalent Compounds

11, 12 Chemical Reactions

13 Giant Molecules

14 The Chemistry and Structure of the Cell

15, 16 Cell Dynamics and the Control of Cellular Activity

17 The Genetic Code: Growth and Replication
18 Cells and Organisms

19 Evolution by Natural Selection
20 Species and Populations

21 Unity and Diversity

22 The Earth: Its Shape, Internal Structure and Composition

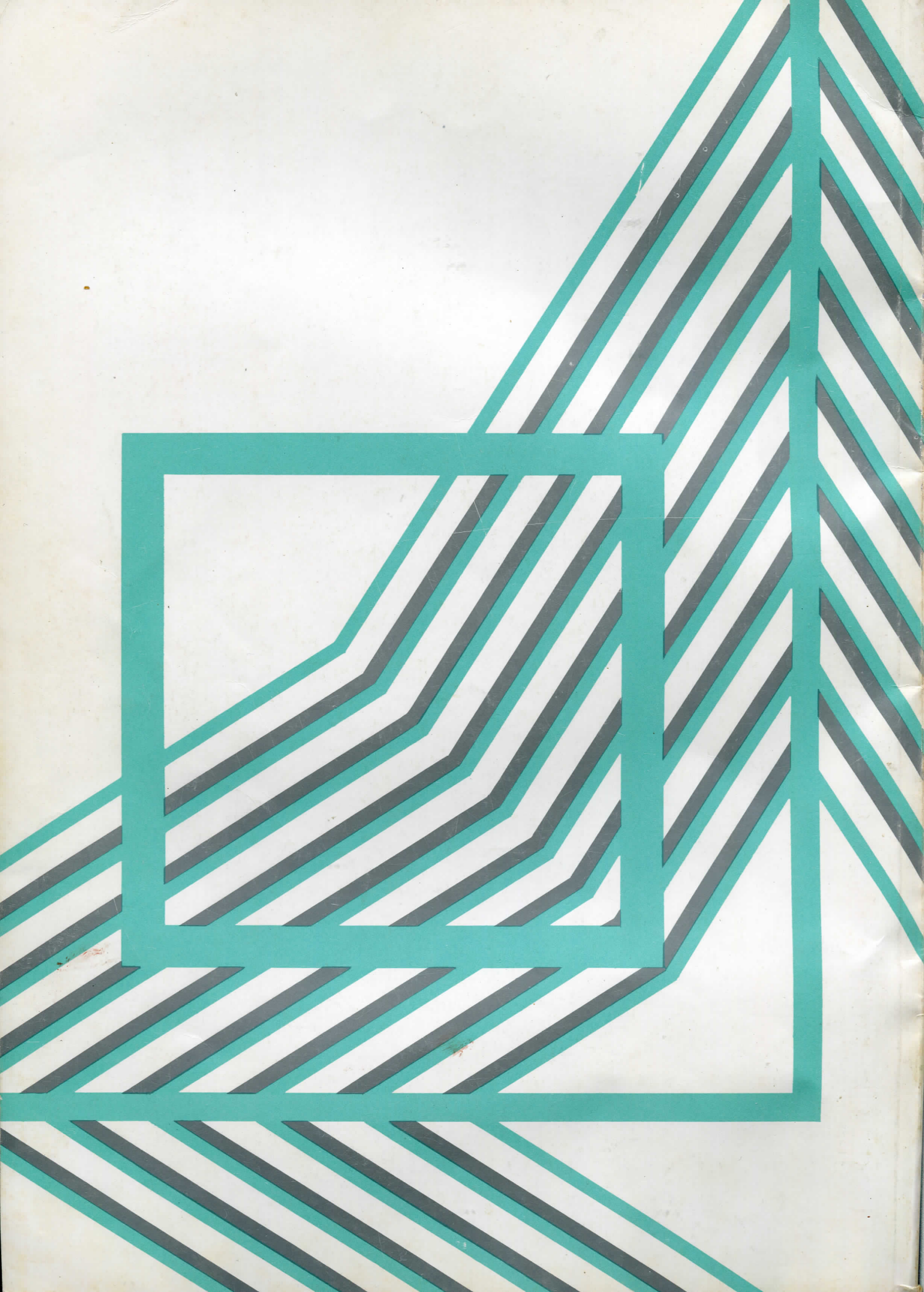
23 The Earth's Magnetic Field

24 Major Features of the Earth's Surface
25 Continental Movement, Sea-floor Spreading and Plate Tectonics

26, 27 Earth History

28 The Wave Nature of Light

29 Quantum Theory
30 Quantum Physics and the Atom

31 The Nucleus of the Atom
32 Elementary Particles

33, 34 Science and Society